

XX
DT
New Mexico

Mature human serum albumin
 Human albumin
 Human serum albumin
 Human serum albumin
 Glycosylated protein
 B lymphocyte stroma
 Human lymphocyte
 Mature form of human
 Recombinant human
 Recombinant human
 Pre human serum albumin
 Human serum albumin
 Human serum albumin
 Human serum albumin
 Cimeric human serum
 Val-Glu-1 human serum
 Human serum albumin
 Human serum albumin
 Glycyl-Glu-1 human serum
 Glycyl-Glu-1 human serum
 Recombinant human serum
 Recombinant human serum
 Prepro-HSA, G-CSE
 HSA-VF1470, 4-HSA
 HSA-VF1470, 4-HSA
 Sequence of mature
 Mature protein of human
 Human serum albumin
 Sequence of human
 Human serum albumin
 Human serum albumin
 Human serum albumin
 Human serum albumin
 Myosin albumin
 Human serum albumin

ALGEBRAISTS

RESULT 1	
ASPECT 18	
10	Antifecia standard, protein: 565 kDa.
11	
12	ASPECT 189:
13	
14	2-NAME-2003 Unlabeled:
15	16 NOV 1993 11:48:00
16	17
17	18 Mature human serum albumin polypeptide.
18	
19	20 Mature human serum albumin polypeptide.
20	
21	22 Mature human serum albumin polypeptide.
22	
23	24 Mature human serum albumin polypeptide.
24	
25	26 Mature human serum albumin polypeptide.
26	
27	28 Mature human serum albumin polypeptide.
28	
29	30 Mature human serum albumin polypeptide.
30	
31	32 Mature human serum albumin polypeptide.
32	
33	34 Mature human serum albumin polypeptide.
34	
35	36 Mature human serum albumin polypeptide.
36	
37	38 Mature human serum albumin polypeptide.
38	
39	40 Mature human serum albumin polypeptide.
40	
41	42 Mature human serum albumin polypeptide.
42	
43	44 Mature human serum albumin polypeptide.
44	
45	46 Mature human serum albumin polypeptide.
46	
47	48 Mature human serum albumin polypeptide.
48	
49	50 Mature human serum albumin polypeptide.
50	
51	52 Mature human serum albumin polypeptide.
52	
53	54 Mature human serum albumin polypeptide.
54	
55	56 Mature human serum albumin polypeptide.
56	
57	58 Mature human serum albumin polypeptide.
58	
59	60 Mature human serum albumin polypeptide.
60	
61	62 Mature human serum albumin polypeptide.
62	
63	64 Mature human serum albumin polypeptide.
64	
65	66 Mature human serum albumin polypeptide.
66	
67	68 Mature human serum albumin polypeptide.
68	
69	70 Mature human serum albumin polypeptide.
70	
71	72 Mature human serum albumin polypeptide.
72	
73	74 Mature human serum albumin polypeptide.
74	
75	76 Mature human serum albumin polypeptide.
76	
77	78 Mature human serum albumin polypeptide.
78	
79	80 Mature human serum albumin polypeptide.
80	
81	82 Mature human serum albumin polypeptide.
82	
83	84 Mature human serum albumin polypeptide.
84	
85	86 Mature human serum albumin polypeptide.
86	
87	88 Mature human serum albumin polypeptide.
88	
89	90 Mature human serum albumin polypeptide.
90	
91	92 Mature human serum albumin polypeptide.
92	
93	94 Mature human serum albumin polypeptide.
94	
95	96 Mature human serum albumin polypeptide.
96	
97	98 Mature human serum albumin polypeptide.
98	
99	100 Mature human serum albumin polypeptide.
100	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

- esp. useful as blood plasma expanders

Disclosure, fig 2, 20pp, English

Mature protein of human serum albumin (see corresp. AAB90C2a).
Used to make new N-terminal fragments which are used as plasma
expanders, or as substitutes for HSA or BSA, in tissue culture
media.

(Updated on 25-MAR-2003 to correct PA field.)

Sequence 585 AA:

Query Match 100.0% Score 3103.00 ID 10 Length 585.
Ref. Seq. Similarity 100.0% Protein No. 9, 58-255.
Matches 585: Conservative C: Mismatches 0: Gaps C:

```

1  DAHSEVAHREKDLSENFKALVLAFAQYLOQCPFEHVKLVNEVTEFANTCVASAE 60
2  DAHSEVAHREKDLSENFKALVLAFAQYLOQCPFEHVKLVNEVTEFANTCVASAE 60
61  NCKSLHTLPQKLCVATRETYGEMACCCAKGPERNECTFQKMDNPNLPRLVPEV 120
121  DMVCTAFHNEETFLKYLVEASRPHTVAPPELLFPAKDYKAAFTCCGAAKACLFP 180
181  KLCLEDEKASSAKORUKCASQKGFERAFKAMAVARLSQRFPAEVSCLYDITLK 240
181  KLCLEDEKASSAKORUKCASQKGFERAFKAMAVARLSQRFPAEVSCLYDITLK 240
241  VHTCCGDLLECCADRADCAKYICENQSSISKKKECEKPLLEKSHCIAEVNDZPA 300
241  VHTCCGDLLECCADRADCAKYICENQSSISKKKECEKPLLEKSHCIAEVNDZPA 300
301  DLPSLAADPVSXKCNKYAEAKOVFLGMFLYEVARHPDYSVLLRLAKTYETLEK 360
301  DLPSLAADPVSXKCNKYAEAKOVFLGMFLYEVARHPDYSVLLRLAKTYETLEK 360
361  CAADPHIEYAKVDFEFPVSEFONLIFKQCELPFEQGEYFONALVRYTKPQVST 420
361  CAADPHIEYAKVDFEFPVSEFONLIFKQCELPFEQGEYFONALVRYTKPQVST 420
421  PTLVNSRNLKYGSKCNKFNENKTYGAEYLSVLAQGLVHENTPVSORVTKCTES 480
421  PTLVNSRNLKYGSKCNKFNENKTYGAEYLSVLAQGLVHENTPVSORVTKCTES 480
481  KSLQVAMDCFAAFVKMCCNADCKMFTFAEVSXKLVAASQAALGL 540
481  KSLQVAMDCFAAFVKMCCNADCKMFTFAEVSXKLVAASQAALGL 540
541  KSLQVAMDCFAAFVKMCCNADCKMFTFAEVSXKLVAASQAALGL 585
541  KSLQVAMDCFAAFVKMCCNADCKMFTFAEVSXKLVAASQAALGL 585

```

RESULT 2

AAB90C2a standard, protein, 585 AA.

AAB90C2a

AAB90C2a

38-OCT-1950 (first entry)

Human serum albumin gene product.

Human serum albumin; HSA-A, Yeast, ds.

Homo sapiens.

CF02117364-A.

585

38-OCT-1950

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

38-OCT-1950 AAB90C2a

X 25-MAR-2003 (updated:
T 15-APR-1991 (first entry)
Y Human serum albumin
X HSA: folding: 68
X Homo sapiens.
S Key
H Region
H Label: A
H Label: B
H Label: C
X JP02227079-A.
X 10-SEP-1990.
X 25-AUG-1989: 892P-0217540
X 25-AUG-1989: 89JF-0217540.
X (TOFU) TONEN CGSF.
X WPI: 1990-317325/42.
X N-PSDB: AAQ0609.
X New human serum albumin fragments - used to bond medicines and for
X stable folding of proteins.
X Claim 1: Fig 8: 24pp: Capalano
X Fragments A-C of HSA are expressed as fusion proteins with the
X signal peptide of E. coli alkaline phosphatase. The fragments are
X referred to as specific proteinase. The C-terminal amino acids
X of the fragments are: A: 1-133, B: 134-214, C: 215-426. The N-terminal truncated
X fragment, C, has good stability in protein folding. The central
X segment, A, has characteristics of both B and C.
X See also AAQ06195-Q06398.
X Updated on 25-MAR-2003 to correct P2 field.
X Updated on 25-MAR-2003 to correct P2 field.
X Updated on 25-MAR-2003 to correct P2 field.
X Sequence 585 AA:
Query Match 100.0% Score 31034 DB 11 Length 585:
Best Local Similarity 100.0% Pred. No. 9, 5k, 255:
Matches 585: Conservative 0, Mismatches 0, Gaps 0.
NY 1 DAHKSEVAHAFKCGEENFKALVIAFAQYLCQTFEDHVKLVNVEVFAKTVADSEAE 60
E 1 DAHKSEVAHAFKCGEENFKALVIAFAQYLCQTFEDHVKLVNVEVFAKTVADSEAE 60
Y 61 NCRKGLHTLPGDKLCTVATREYTGEMACCKKQEPFNCFLQKHCHNPNLPVPEV 120
D 61 NCRKGLHTLPGDKLCTVATREYTGEMACCKKQEPFNCFLQKHCHNPNLPVPEV 120
Y 121 DVNCTAFHNEETLKKYLYVEIARRHPYVAPVPELFFAKRYKANTTECCQADKAGCLLP 180
D 121 DVNCTAFHNEETLKKYLYVEIARRHPYVAPVPELFFAKRYKANTTECCQADKAGCLLP 180
Y 181 KLDELROESKASSAKORLKCAISQKFRERAFKAMAVARLSQRFPAEFAEVSKLVTDTK 240
D 181 KLDELROESKASSAKORLKCAISQKFRERAFKAMAVARLSQRFPAEFAEVSKLVTDTK 240
Y 241 VHTCCGMLLEADGADALANYIGENQDISSSKMKKEZCEKTLPRKSHAEVNEEMPA 300
D 241 VHTCCGMLLEADGADALANYIGENQDISSSKMKKEZCEKTLPRKSHAEVNEEMPA 300

QY 301 RASFSLAOFVESKIVCYRYAEAKVFLQMFYVARRHFDYSVGLLELAKTYETILEMC 360
DY 301 RASFSLAOFVESKIVCYRYAEAKVFLQMFYVARRHFDYSVGLLELAKTYETILEMC 360
QY 361 GAAGDHECHAKKDKNEKHSFEPNLIKQNFSPQLQFVKFONALNRYTKYFOUST 420
DY 361 GAAGDHECHAKKDKNEKHSFEPNLIKQNFSPQLQFVKFONALNRYTKYFOUST 420
QY 421 PTLVEYSKLVGVGSKKCHHFEAKRMPCAEDYLSVYNOLQVLREKTEVSRAVTKCTES 480
DY 421 PTLVEYSKLVGVGSKKCHHFEAKRMPCAEDYLSVYNOLQVLREKTEVSRAVTKCTES 480
QY 481 DVNRPFSALREYHRYVYKFNATETTFHAGICTLSEKERQIKKQALVELVGRPKAT 540
DY 481 DVNRPFSALREYHRYVYKFNATETTFHAGICTLSEKERQIKKQALVELVGRPKAT 540
QY 541 YEGCKAYMCOFAFVYKCKADKQKCECFENESKZKLVAAACAALGL 595
DY 541 YEGCKAYMCOFAFVYKCKADKQKCECFENESKZKLVAAACAALGL 595
RESULT 4
AA09030:
TC AA090301 Standard: Protein: 585 AA.
XX AA090301:
XX 25-MAR-2003 (updated:
DT 25-MAR-1990 (first entry)
XX Human serum albumin
XX Seron albumin, HSA, aspartyl protease-3, YAP3p:
XX Saccharomyces cerevisiae.
XX Homo sapiens.
XX W0292387 A:
XX DB 05F-1993.
XX 01-MAR-1993, 95MD 2B03434
XX 05-MAR-1994, 440B 0104270.
XX TABLE: DELTA BIOTECHNOLOGY LTD.
XX KerryWilliams SM, Gilbert SC,
XX NDI: 1896-32057-41:
XX N-PSDB: AAQ0609.
XX Yeast with induced levels of aspartyl protease 3 proteolytic
XX activity used to succinate human albumin without proteolysis of the 45
XX KD fragment
XX Example 1: Page 26-28, 50pp, English.
XX The cDNA given in AA09030, which encodes HSA (AA09030), was subjected
XX to site-directed mutagenesis to investigate the role of
XX endoproteases in the generation of a 45 kDa albumin fragment obd-
XX in the degradation of HSA. The cDNA was mutated to replace the
XX K439A (K439V, 4509A) and K440A (K440V, 4510A) mutations were of
XX this type. The mutations resulted in a 45 kDa albumin fragment of
XX proteolytic cleavage, allowing increased proteolysis of recombinant HSA.
XX Updated on 25-MAR-2003 to correct P1 field.
XX Sequence 585 AA.
Query Match 100.0% Score 3103 DB 11 Length 585:
Best Local Similarity 100.0% Pred. No. 9, 5k, 255:
Matches 585: Conservative 0, Mismatches 0, Indels 0, Gaps 0.

/note: *optionally acetylated, and claimed under claim 56"

WC00020940-A1

11-APR-2000

31-OCT-1999: 93NC-US22503

32-OCT-1998: 96US-0102738

02-OCT-1998: 96US-0165841

02-OCT-1998: 96US-0165926

11-JAN-1999: 96US-0155332

USCH-1 SCHEMIA TECHNOLOGIES INC.

Bar-Or E, Loo E, Winkler JV.

WPI: 2000-033447/26

New method for the continuous detection of ischemic states comprises detecting and quantifying the existence of an alteration of the serum protein albumin.

Disclosure: Page 97-100; 105pp; English

The present sequence represents human albumin protein. The specification describes a method for the continuous detection of ischemic states. The method comprises detecting and quantifying the existence of an alteration of the serum protein albumin. The method comprises contacting a serum sample with a reagent that binds to the N-terminus of naturally occurring human albumin, to form a mixture containing bound metal ions and unbound metal ions, and then determining the amount of metal ions bound to the albumin N-terminus. The amount of bound metal ions is correlated to a known value to determine the occurrence or non-occurrence of an ischemic event. The methods are useful for detection of ischemic states. The methods are also useful for distinguishing periparturient ischemia from ischemia caused by, amongst other things, myocardial infarction and progressive coronary artery disease.

Sequence 595 AA:

Query Match 100.0% Score 3103; DB 21; Length 585.
Best Local Similarity 100.0%; Pred. No. 9, 5e-255;
Matches 585; Conservative C; Mismatches 0; Indels 0; Gaps 0;

1 DAKSEVAFHFXDGEENFKALVLAFAQYLCQQCFEDHVKLVNVEFAKCVVAERSAE 60
1 DAKSEVAFHFXDGEENFKALVLAFAQYLCQQCFEDHVKLVNVEFAKCVVAERSAE 60
61 NCKDSHFTLQKLTVALBETYSGMADCAKQEPNECFQHKEDNPNLPVAREV 120
61 NCKDSHFTLQKLTVALBETYSGMADCAKQEPNECFQHKEDNPNLPVAREV 120
121 DWCTAFHDMETFLKPYVEIARHPHYVFAPELLFAKPYKAAFTCCQAADYAAQLDF 180
121 DWCTAFHDMETFLKPYVEIARHPHYVFAPELLFAKPYKAAFTCCQAADYAAQLDF 180
181 KLDELNDPMASAKQLKCLASLQKQFRAFPAKAVARLSORTFAFPAFVKNLYTDLY 240
181 KLDELNDPMASAKQLKCLASLQKQFRAFPAKAVARLSORTFAFPAFVKNLYTDLY 240
241 VHTCCCHGCLLEADDAKALYICNDQSSISKLKCCCEPFLKSHCAEVENKXPA 300
241 VHTCCCHGCLLEADDAKALYICNDQSSISKLKCCCEPFLKSHCAEVENKXPA 300
301 DLFSLAAPFESKQYCKVAEKVPLQMLFVEYARRHPYVSYVLLRLAKYETLDEVC 360
301 DLFSLAAPFESKQYCKVAEKVPLQMLFVEYARRHPYVSYVLLRLAKYETLDEVC 360
361 CAADPHCEYAFKPEKPLNVEEPQNLKNCDELFCQGLGEYKFNALDVHVKAPVYST 420

361 CAADPHCEYAFKPEKPLNVEEPQNLKNCDELFCQGLGEYKFNALDVHVKAPVYST 420
401 FTVVPMHMLKGVSKQVWFFFAKPYKAAFTCCQAADYAAQLDF 460
401 FTVVPMHMLKGVSKQVWFFFAKPYKAAFTCCQAADYAAQLDF 460
421 FTVVPMHMLKGVSKQVWFFFAKPYKAAFTCCQAADYAAQLDF 480
421 FTVVPMHMLKGVSKQVWFFFAKPYKAAFTCCQAADYAAQLDF 480
441 LNKQSTPSHVEYVYKPEKATETTHANICTLSEKPEQIKYKTKALVEYVHKNPAT 540
441 LNKQSTPSHVEYVYKPEKATETTHANICTLSEKPEQIKYKTKALVEYVHKNPAT 540
461 LNKQSTPSHVEYVYKPEKATETTHANICTLSEKPEQIKYKTKALVEYVHKNPAT 540
461 LNKQSTPSHVEYVYKPEKATETTHANICTLSEKPEQIKYKTKALVEYVHKNPAT 540
541 NQCAKAVMCDFAATVYKQKADKNETTFAEKQKVLVAASCAALGI 595
541 NQCAKAVMCDFAATVYKQKADKNETTFAEKQKVLVAASCAALGI 595

RESULT 7
AAVS948
AAVS948 standard; Protein: 595 AA
AC AAVS948
AC AAVS948
UT 25-JUL-2003 (first entry)
XX Yeast codon-biased recombinant human serum albumin protein.
XX Recombinant human serum albumin, HSA; Yeast codon bias; Host cell.
XX over-lapping oligonucleotide expression vector.
XX Homo sapiens.
XX Synthesis.
XX CNL23303A.
XX 22 DEC-1999.
XX 17-JUN-1998: 98CN-0102506.
XX 17-JUN-1998: 98CN-0102536.
XX (HA)-1-1 HAIT BIOENGINEERING CO LTD
XX LI S, Lu D:
XX WPI: 2000-033447/26;
XX N-PCR: AAA10091.
XX Process for preparing recombinant human serum albumin - which comprises
XX Yeast biased sex codons
XX Disclosure: Fig 1: 44pp; Chinese.
XX The method relates to a method of recombinantly producing human serum
XX albumin (HSA) in yeast by altering the coding sequence of HSA to
XX comprise a yeast codon bias. The complete HSA gene (AAA10091) was
XX generated as three synthetic fragments (AAA10092-10094) joined by
XX recombinant DNA technology. Each HSA fragment was synthesised from
XX overlapping oligonucleotide fragments that were extended. This sequence
XX represents the complete sequence of the HSA encoded by the human gene
XX with a yeast codon bias. The invention also covers a recombinant
XX expression vector, yeast host cells carrying the recombinant expression
XX vector and the process for producing human serum albumin in the yeast
XX host cells, especially in secretory mode.
XX Sequence 595 AA:
Query Match 100.0% Score 3103; DB 21; Length 585.
Best Local Similarity 100.0%; Pred. No. 9, 5e-255;
Matches 585; Conservative C; Mismatches 0; Indels 0; Gaps 0;

1 DAKSEVAFHFXDGEENFKALVLAFAQYLCQQCFEDHVKLVNVEFAKCVVAERSAE 60
1 DAKSEVAFHFXDGEENFKALVLAFAQYLCQQCFEDHVKLVNVEFAKCVVAERSAE 60
61 NCKDSHFTLQKLTVALBETYSGMADCAKQEPNECFQHKEDNPNLPVAREV 120
61 NCKDSHFTLQKLTVALBETYSGMADCAKQEPNECFQHKEDNPNLPVAREV 120
121 DWCTAFHDMETFLKPYVEIARHPHYVFAPELLFAKPYKAAFTCCQAADYAAQLDF 180
121 DWCTAFHDMETFLKPYVEIARHPHYVFAPELLFAKPYKAAFTCCQAADYAAQLDF 180
181 KLDELNDPMASAKQLKCLASLQKQFRAFPAKAVARLSORTFAFPAFVKNLYTDLY 240
181 KLDELNDPMASAKQLKCLASLQKQFRAFPAKAVARLSORTFAFPAFVKNLYTDLY 240
241 VHTCCCHGCLLEADDAKALYICNDQSSISKLKCCCEPFLKSHCAEVENKXPA 300
241 VHTCCCHGCLLEADDAKALYICNDQSSISKLKCCCEPFLKSHCAEVENKXPA 300
301 DLFSLAAPFESKQYCKVAEKVPLQMLFVEYARRHPYVSYVLLRLAKYETLDEVC 360
301 DLFSLAAPFESKQYCKVAEKVPLQMLFVEYARRHPYVSYVLLRLAKYETLDEVC 360
361 CAADPHCEYAFKPEKPLNVEEPQNLKNCDELFCQGLGEYKFNALDVHVKAPVYST 420


```

CITY, KING OF RUSS-A
STATE, RUSSIA
COUNTRY, RUSSIA
ZIP, 19406-1112
CONTACT, FRANKLIN FORM
MEDIUM TYPE, FIDITY disk
COMPUTER, IBM PC COMPATIBLE
OPERATING SYSTEM, MS-DOS
SOFTWARE, MICROSOFT WORD 6.0
CURRENT APPLICATION DATA,
PUBLICATION NUMBER, 98/702,572
FILING DATE, 1-NOV-1996
CLASSIFICATION, 415
PRIORITY APPLICATION DATA,
APPLICATION NUMBER, WC 95/23557
FILING DATE, 1-MAR-1994
APPLICATION NUMBER, GB 9404275.2
FILING DATE, 5-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME, Naomi Biswar
REGISTRATION NUMBER, 18,164
REFERENCE/EXAMINER NUMBER, GE0314 US
TELEPHONE, 610/818-4294
TELEFAX, 610/818/4421
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS
LENGTH, 586 amino acids
TYPE, amino acid
TOPOLOGY, linear
MOLECULE TYPE, Protein
CS-GB-702-572.2

Query March 1997, Score 15.3, DB 2, Length 585.
Best local similarity 190.08, P-Val No. 9,30-287, Indels 0, Gaps 3.
Matches 585, Conservative 0, Mismatches 0.

QY 1 DAKSEVAHRKDSGKGNFKALYLAPAOYLQCCPFEDHVKLYVNEVTEFAKTCVAESSE 60
DE 1 DAKSEVAHRKDSGKGNFKALYLAPAOYLQCCPFEDHVKLYVNEVTEFAKTCVAES 60
DY 1 DKSRLTILYDKMCKVATLLETGSDACCKAKGSEKSNSTCFQHKDQDNPLPRVPEV 120
DB 51 DKSRLTILYDKMCKVATLLETGSDACCKAKGSEKSNSTCFQHKDQDNPLPRVPEV 120
QY 131 DVMCTAFHNETLKKLYVEIARRHVPYAEPLLFPKPKYKAAFTCCCAAKKALCLP 180
DB 131 DVMCTAFHNETLKKLYVEIARRHVPYAEPLLFPKPKYKAAFTCCCAAKKALCLP 180
QY 151 KLELRGSKASAZQKCKASIOKFFERAFKAMAVRISGRPKKAEFAVSKVTDLK 240
DB 151 KLELRGSKASAZQKCKASIOKFFERAFKAMAVRISGRPKKAEFAVSKVTDLK 240
QY 241 VHTCECHDILFQCDPAFLAKYICENQDSISKVCECKEPLLEKSHCTASVENVK 300
DB 241 VHTCECHDILFQCDPAFLAKYICENQDSISKVCECKEPLLEKSHCTASVENVK 300
QY 301 DLPSEKAEFVSKQVKKVAGVQVLFQNVYVFAHPDVSVDLLKLAKTVETILEK 360
DB 301 DLPSEKAEFVSKQVKKVAGVQVLFQNVYVFAHPDVSVDLLKLAKTVETILEK 360

```

Qy	361	QALAPHECHAKVPEDEPRVUSPQMLKONTEPEFDESEKFKONALLVYTKVPOVST	420
Dc	362	QALAPHECHAKVPEDEPRVUSPQMLKONTEPEFDESEKFKONALLVYTKVPOVST	420
Qy	421	FTJLVEKPLKAKVCKKHEKAKPVANDVLSVNLQCLVHEKTVSDRVTKCTES	480
Dc	422	FTJLVEKPLKAKVCKKHEKAKPVANDVLSVNLQCLVHEKTVSDRVTKCTES	480
Qy	481	LVMPKATVSAALVYVYVTKVPEKHAETTFKACITITISEKEROYKCATVELVYKHKPKAT	540
Dc	482	LVMPKATVSAALVYVYVTKVPEKHAETTFKACITITISEKEROYKCATVELVYKHKPKAT	540

SEQUENCE CHARACTERISTICS:

LENGTH: 609 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-433-037-4

Query Match: 99.9%; Score 3099; DB 1; Length 609;

Best Local Similarity 99.8%; Pred. No. 2.4e-286;

Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 DAHSEVAHREFKDLGEENFKALVLIAPAOYLQCCPREDHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAHSEVAHREFKDLGEENFKALVLIAPAOYLQCCPREDHVKLVNEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEBNECFQHKDDNPNJFRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEBNECFQHKDDNPNJFRLVRPEV 144
QY 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFPAKYKAAFTCCCOADKAACLIP 180
DB 145 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFPAKYKAAFTCCCOADKAACLIP 204
QY 181 KLDELREGKASSAKORLKCASLOKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240
DB 205 KLDELREGKASSAKORLKCASLOKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPJJLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPJJLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNAYAEAKDVFLGMFLYBYARRHPDYSVVLRLAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNAYAEAKDVFLGMFLYBYARRHPDYSVVLRLAKTYETTLK 384
QY 361 CAADHPHECYAKVFDEPKPLVEEPQNLIKQNCBLFEQLGYKFNALLVRYTKKVPQVST 420
DB 385 CAADHPHECYAKVFDEPKPLVEEPQNLIKQNCBLFEQLGYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480
DB 445 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 504
QY 481 LVNRRPFCFSALEVDENVYVPKEFNAETFTFHADCTLSEKERQIKKQTALVELVKHKPKAT 540
DB 505 LVNRRPFCFSALEVDENVYVPKEFNAETFTFHADCTLSEKERQIKKQTALVELVKHKPKAT 564
QY 541 KEQLKAVMDQFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
DB 565 KEQLKAVMDQFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 609

```

RESULT 10

US-08-897-956A-2

Sequence 2, Application US/08897956A

Patent No. 6423512

GENERAL INFORMATION:

APPLICANT: Mary Ellen Digan

APPLICANT: Philip Lake

APPLICANT: Hermann Gram

TITLE OF INVENTION: Fusion Polypeptides

FILE REFERENCE: 600-7244/CPA

CURRENT APPLICATION NUMBER: US/08/897.956A

CURRENT FILING DATE: 1997-07-21

PRIOR APPLICATION NUMBER: 60/022.689

PRIOR FILING DATE: 1996-07-26

NUMBER OF SEQ ID NOS: 38

SOFTWARE: Fast-Seq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 609

TYPE: PRT

ORGANISM: Homo Sapiens

US-08-897-956A-2

Query Match: 99.9%; Score 3099; DB 4; Length 609;

Best Local Similarity 99.8%; Pred. No. 2.4e-286;

Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 DAHSEVAHREFKDLGEENFKALVLIAPAOYLQCCPREDHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAHSEVAHREFKDLGEENFKALVLIAPAOYLQCCPREDHVKLVNEVTEFAKTCVADESAE 94
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEBNECFQHKDDNPNJFRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEBNECFQHKDDNPNJFRLVRPEV 144
QY 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFPAKYKAAFTCCCOADKAACLIP 180
DB 145 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFPAKYKAAFTCCCOADKAACLIP 204
QY 181 KLDELREGKASSAKORLKCASLOKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240
DB 205 KLDELREGKASSAKORLKCASLOKGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPJJLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPJJLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNAYAEAKDVFLGMFLYBYARRHPDYSVVLRLAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNAYAEAKDVFLGMFLYBYARRHPDYSVVLRLAKTYETTLK 384
QY 361 CAADHPHECYAKVFDEPKPLVEEPQNLIKQNCBLFEQLGYKFNALLVRYTKKVPQVST 420
DB 385 CAADHPHECYAKVFDEPKPLVEEPQNLIKQNCBLFEQLGYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480
DB 445 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 504
QY 481 LVNRRPFCFSALEVDENVYVPKEFNAETFTFHADCTLSEKERQIKKQTALVELVKHKPKAT 540
DB 505 LVNRRPFCFSALEVDENVYVPKEFNAETFTFHADCTLSEKERQIKKQTALVELVKHKPKAT 564
QY 541 KEQLKAVMDQFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
DB 565 KEQLKAVMDQFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 609

```

RESULT 11

PCT-US95-04375-3

Sequence 3, Application PC/TUS9504075

GENERAL INFORMATION:

APPLICANT: AMGEN INC.

TITLE OF INVENTION: Afamir: A Human Serum Albumin-Like

TITLE OF INVENTION: Protein

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amgen Center, Patent Operations/RR

STREET: 1840 DeHavilland Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: U.S.

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

FILING DATE:

CLASSIFICATION:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 609 amino acids

TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MCLEUCULE TYPE: protein
PCT-US95-04075-3

Query Match 99.9%; Score 3099; DB 5; Length 609;
Best Local Similarity 99.8%; Pred. No. 2.4e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRRFKDLGEENFKALVLIAPAFAYLQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHKEVAHRRFKDLGEENFKALVLIAPAFAYLQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDNPRLVLRREV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDNPRLVLRREV 144

QY 121 DVMCTAFHNDNEETFLKKYLSEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 180
DB 145 DVMCTAFHNDNEETFLKKYLSEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 204

QY 181 KLDELRDGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFFPKAEPAEVSUKLVTDLTK 240
DB 205 KLDELRDGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFFPKAEPAEVSUKLVTDLTK 264

QY 241 VHTCCHGDLLECADRADLAKYICENODSISSKLKECCCKPDLLEKSHCAEVENDEMPA 300
DB 265 VHTCCHGDLLECADRADLAKYICENODSISSKLKECCCKPDLLEKSHCAEVENDEMPA 324

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGNFLYEYARRHPDYSVLLRLAKTYETTLEKC 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFLGNFLYEYARRHPDYSVLLRLAKTYETTLEKC 384

QY 361 CAAADPHCYAKVDFEFPKLPVEEPONLTKONCELPQGLGEYKFNQALLVRYTKVPQVST 420
DB 385 CAAADPHCYAKVDFEFPKLPVEEPONLTKONCELPQGLGEYKFNQALLVRYTKVPQVST 444

QY 421 PTLVEVSRNLGVGSKCKGHPKAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGVGSKCKGHPKAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALVEVDYTPVPKEFNAETFTPHADICTLSEKERQIKKQATLVELVHKHPKAT 540
DB 505 LVNRRPCFSALVEVDYTPVPKEFNAETFTPHADICTLSEKERQIKKQATLVELVHKHPKAT 564

QY 541 KEQLKAVMDDFAAVFEKCKCKADKTCFAEKGKLVAAASQAALGL 585
DB 565 KEQLKAVMDDFAAVFEKCKCKADKTCFAEKGKLVAAASQAALGL 609

RESULT 12
US-08-897-956A-3
Sequence 3, Application US/08897956A
Patent No. 6423512
GENERAL INFORMATION:
APPLICANT: Mary Ellen Digan
APPLICANT: Philip Lake
APPLICANT: Hermann Gram
TITLE OF INVENTION: Fusion Polypeptides
FILE REFERENCE: 609-7244/CPA
CURRENT APPLICATION NUMBER: US/08/897,956A
CURRENT FILING DATE: 1997-07-21
PRIOR APPLICATION NUMBER: 60/022,689
PRIOR FILING DATE: 1996-07-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 978
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Fusion polypeptide

US-08-897-956A-3

Query Match 99.7%; Score 3095; DB 4; Length 978;
Best Local Similarity 99.8%; Pred. No. 2.1e-285;
Matches 583; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRRFKDLGEENFKALVLIAPAFAYLQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 212 DAHKEVAHRRFKDLGEENFKALVLIAPAFAYLQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 271

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDNPRLVLRREV 120
DB 272 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDNPRLVLRREV 331

QY 121 DVMCTAFHNDNEETFLKKYLSEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 180
DB 332 DVMCTAFHNDNEETFLKKYLSEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 391

QY 191 KLDELRDGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFFPKAEPAEVSUKLVTDLTK 240
DB 392 KLDELRDGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFFPKAEPAEVSUKLVTDLTK 451

QY 241 VHTCCHGDLLECADRADLAKYICENODSISSKLKECCCKPDLLEKSHCAEVENDEMPA 300
DB 452 VHTCCHGDLLECADRADLAKYICENODSISSKLKECCCKPDLLEKSHCAEVENDEMPA 511

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGNFLYEYARRHPDYSVLLRLAKTYETTLEKC 360
DB 512 DLPSLAADFVESKDVCKNYAEAKDVFLGNFLYEYARRHPDYSVLLRLAKTYETTLEKC 571

QY 361 CAAADPHCYAKVDFEFPKLPVEEPONLTKONCELPQGLGEYKFNQALLVRYTKVPQVST 420
DB 572 CAAADPHCYAKVDFEFPKLPVEEPONLTKONCELPQGLGEYKFNQALLVRYTKVPQVST 631

QY 421 PTLVEVSRNLGVGSKCKGHPKAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 632 PTLVEVSRNLGVGSKCKGHPKAKMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 691

QY 481 LVNRRPCFSALVEVDYTPVPKEFNAETFTPHADICTLSEKERQIKKQATLVELVHKHPKAT 540
DB 692 LVNRRPCFSALVEVDYTPVPKEFNAETFTPHADICTLSEKERQIKKQATLVELVHKHPKAT 751

QY 541 KEQLKAVMDDFAAVFEKCKCKADKTCFAEKGKLVAAASQAALGL 584
DB 752 KEQLKAVMDDFAAVFEKCKCKADKTCFAEKGKLVAAASQAALGL 795

RESULT 13
US-08-448-196A-3
Sequence 3, Application US/08448196A
Patent No. 5780594
GENERAL INFORMATION:
APPLICANT: CARTER, DANIEL C.
TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
TITLE OF INVENTION: RELATED PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: NASA
STREET: MARSHALL SPACE FLIGHT CENTER
CITY: HUNTSVILLE
STATE: ALABAMA
COUNTRY: USA
ZIP: 35812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,196A
FILING DATE: 23-MAY-1995
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
 NAME: BROAD JR., ROBERT L.
 REGISTRATION NUMBER: 18,757
 REFERENCE/DOCKET NUMBER: XX/WFS-28432-2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 205-544-0021
 TELEFAX: 205-544-0258
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 585 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEetical: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: N-terminal
 US-08-448-196A-3

Query Match 99.7%; Score 3093; DB 1; Length 585;
 Best Local Similarity 99.7%; Pred. No. 8.3e-286;
 Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DAHSEVAHREFKDLGGEENFKALVLIIFAQYLOQCPPEHVKLVNEVTEFAKTCVADESAAE 60
 Db 1 DAHSEVAHREFKDLGGEENFKALVLIIFAQYLOQCPPEHVKLVNEVTEFAKTCVADESAAE 60

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120

Qy 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180
 Db 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180

Qy 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQFPKAEFAEVSKLVTDLTK 240
 Db 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQFPKAEFAEVSKLVTDLTK 240

Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDMPA 300
 Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDMPA 300

Qy 301 DPLSLAADFVESKDVCKNYAEAKDVFGLMFLVEYARRHPDYSVLLRLAKTYETTTLEKC 360
 Db 301 DPLSLAADFVESKDVCKNYAEAKDVFGLMFLVEYARRHPDYSVLLRLAKTYETTTLEKC 360

Qy 361 CAADHPHECYAKVDFDEPKLVBEPPQLIKONCELFQKQGEYKFNALLVRYTKVPQVST 420
 Db 361 CAADHPHECYAKVDFDEPKLVBEPPQLIKONCELFQKQGEYKFNALLVRYTKVPQVST 420

Qy 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCCTES 480
 Db 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCCTES 480

Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540
 Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540

Qy 541 KEQLKAVMDDFAAFEVKCKCKADDKETCFABEGKKLVAASQAALGL 585
 Db 541 KEQLKAVMDDFAAFEVKCKCKADDKETCFABEGKKLVAASQAALGL 585

RESULT 14

US-08-984-176-1
 Sequence 1, Application US/08984176
 Patent No. 5948609
 GENERAL INFORMATION:
 APPLICANT: CARTER, DANIEL C
 APPLICANT: HO, JOSEPH X
 APPLICANT: RUKER, FLORIAN
 TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
 TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER

FILE REFERENCE: 08/984,176
 CURRENT APPLICATION NUMBER: US/08/984,176
 CURRENT FILING DATE: 1997-12-03
 NUMBER OF SEQ ID NOS: 1
 SOFTWARE: Patentit Ver. 2.0
 SEQ ID NO 1
 LENGTH: 585
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-08-984-176-1

Query Match 99.7%; Score 3093; DB 2; Length 585;
 Best Local Similarity 99.7%; Pred. No. 8.3e-286;
 Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DAHSEVAHREFKDLGGEENFKALVLIIFAQYLOQCPPEHVKLVNEVTEFAKTCVADESAAE 60
 Db 1 DAHSEVAHREFKDLGGEENFKALVLIIFAQYLOQCPPEHVKLVNEVTEFAKTCVADESAAE 60

Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120

Qy 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180
 Db 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180

Qy 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQFPKAEFAEVSKLVTDLTK 240
 Db 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQFPKAEFAEVSKLVTDLTK 240

Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDMPA 300
 Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDMPA 300

Qy 301 DPLSLAADFVESKDVCKNYAEAKDVFGLMFLVEYARRHPDYSVLLRLAKTYETTTLEKC 360
 Db 301 DPLSLAADFVESKDVCKNYAEAKDVFGLMFLVEYARRHPDYSVLLRLAKTYETTTLEKC 360

Qy 361 CAADHPHECYAKVDFDEPKLVBEPPQLIKONCELFQKQGEYKFNALLVRYTKVPQVST 420
 Db 361 CAADHPHECYAKVDFDEPKLVBEPPQLIKONCELFQKQGEYKFNALLVRYTKVPQVST 420

Qy 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCCTES 480
 Db 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVNLQCLVLEKTPVSDRVTKCCCTES 480

Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540
 Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540

Qy 541 KEQLKAVMDDFAAFEVKCKCKADDKETCFABEGKKLVAASQAALGL 585
 Db 541 KEQLKAVMDDFAAFEVKCKCKADDKETCFABEGKKLVAASQAALGL 585

RESULT 15

US-08-448-196A-5
 Sequence 5, Application US/08448196A
 Patent No. 5780594
 GENERAL INFORMATION:
 APPLICANT: CARTER, DANIEL C
 TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
 TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
 TITLE OF INVENTION: RELATED PROTEINS
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NASA
 STREET: MARSHALL SPACE FLIGHT CENTER
 CITY: HUNTSVILLE
 STATE: ALABAMA
 COUNTRY: USA
 ZIP: 35812

Job time : 31 secs

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,196A
FILING DATE: 23-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: BROAD JR., ROBERT L.
REGISTRATION NUMBER: 18,757
REFERENCE/DOCKET NUMBER: XX/MFS-28432-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 205-544-0021
TELEFAX: 205-544-0258
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-termina
US-08-448-196A-5

Query Match 79.2%; Score 2458.5; DB 1; Length 583;
Best Local Similarity 75.8%; Pred. NC. 2.1e-225;
Matches 442; Conservative 70; Mismatches 70; Indels :; Gaps 1;
QY 1 DAHKSEVAHRFKXOLGEENFKALVLIAPAOYLQCCPEDHDVHVKLVNEVTEFAKTCVADESAB 60
DB 1 DTHKSEIAHRFNDLGEKFKGLVLFVPSQYLQCCPEDHDVHVKLVNEVTEFAKKCADESAB 60
QY 61 NCDKSLHTJFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDRNLPRLVRPEV 120
DB 61 NCDKSLHTJFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDRNLPRLVRPEV 120
QY 121 DVNCTAFHDNEETFLKXYLYEIAARRHPYFYVAPELLFFAKYKAAFTCCOAAADKAACLLP 180
DB 120 DAQCAAFCEZDFDKFLGYLYEVARRRHPYFYGPPELLFHABEYKADFTCECCPADDKLALIP 179
QY 181 KLDELDRDEGKASSAKQRLKASLOKFGERAFKAMAVARLSQRFPKAEFAEVSXKLVTDJTK 240
DB 180 KLDALKERILLSSAKERLKCSSFQNGERAVKAVSARLSQKFPKADFAEVSXKLVTDJTK 239
QY 241 VHTCCCHGDLLECAADDRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 300
DB 240 VHKCCCHGDLLECAADDRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 299
QY 301 DLPSLAADFVESKDVKKNYAEAKDVFLGMFLEYEARHPDYSVVLLRLAKTYETLEKC 360
DB 300 DIPALAADFAEDKEIGKYKDAKDVFLGTFLEYESRRHPDYSVSLLR:AKTYEATLEKC 359
QY 361 CAADPHCEYAKVDEFKPLVEEPQNI:KQNCLEFQELGEYKFQNALLVRYTKVPQVST 420
DB 360 CAEADPPACYRTVPDQFTPLVEEPKSLVKKNCDLFEVGEYDFQNALIVRYTKKAPQVST 419
QY 421 PTLVEVRNLCKVSKCKKPEAKRMPCAEDYLSVNLQ:LVLEHCTPVSDRVTKCCTES 480
DB 420 PTLVEIGRTLKVGSRCKLPESERLPCSENHLALANRLCVLHKEKTPVSEKIKTCCTDS 479
QY 481 LVNRRPCPSALEVDETYVPKEFNAETTFHADICTLSEKERQIKKQ:ALVELVKKHKPKAT 540
DB 480 LAERRPCCPSALELEDCG:PVKPEFKAETTFHADICTLPEDEKQ:KKQSALAE:LVKKEKPKAT 539
QY 541 KEQLKAVNDFFAAVFVEKCKKADDKETCFABEESKLVVAASQAAL 583
DB 540 KEQLKTVLGNFSAFVAKCCGREDKEACFAEEGPKLVASSQLAL 582

GenCore version 5.2.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

DM protein - protein search, using sw model

Run on: October 27, 2003, 15:21:13 ; Search time 72 Seconds
(without alignments)

1360.618 Million cell updates/sec

Title: US-09-833-117-18

Perfect score: 3103

Sequence: 1 DAHSEVAHRKDK/GEENFK.....TCFAEESKLVASQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US05_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_FJBOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US0A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	585	10	US-09-929-552-2
2	3103	100.0	585	11	US-09-932-63-445
3	3103	100.0	585	11	US-09-984-010-26
4	3103	100.0	585	11	US-09-833-041-18
5	3103	100.0	585	12	US-10-153-604A-5
6	3103	100.0	585	12	US-09-833-117-18
7	3103	100.0	585	12	US-10-319-263-1
8	3103	100.0	585	12	US-10-319-263-2
9	3103	100.0	585	12	US-10-414-469-1
10	3103	100.0	585	12	US-10-414-469-2
11	3103	100.0	585	12	US-09-932-322-445
12	3103	100.0	585	12	US-10-413-831-1
13	3103	100.0	585	12	US-10-413-831-2
14	3103	100.0	585	14	US-10-153-064-5
15	3103	100.0	634	11	US-09-984-010-7

16	3103	100.0	609	11	US-09-939-039-370
17	3103	100.0	609	12	US-10-153-604A-7
18	3103	100.0	609	12	US-10-365-623-23
19	3103	100.0	609	14	US-10-153-064-7
20	3103	100.0	610	10	US-09-984-186-2
21	3103	100.0	610	15	US-10-237-667-2
22	3103	100.0	610	15	US-10-237-708-2
23	3103	100.0	610	15	US-10-237-866-2
24	3103	100.0	610	15	US-10-237-871-2
25	3103	100.0	610	15	US-10-237-624-2
26	3103	100.0	651	12	US-10-153-604A-133
27	3103	100.0	651	14	US-10-153-064-133
28	3103	100.0	652	12	US-10-153-604A-132
29	3103	100.0	652	14	US-10-153-064-132
30	3103	100.0	653	12	US-10-153-604A-131
31	3103	100.0	653	14	US-10-153-064-131
32	3103	100.0	656	12	US-10-153-604A-130
33	3103	100.0	656	14	US-10-153-064-130
34	3103	100.0	676	12	US-10-153-604A-129
35	3103	100.0	676	12	US-10-153-604A-129
36	3103	100.0	676	14	US-10-153-064-129
37	3103	100.0	677	12	US-10-153-604A-125
38	3103	100.0	677	14	US-10-153-064-125
39	3103	100.0	680	12	US-10-153-604A-123
40	3103	100.0	680	14	US-10-153-064-123
41	3103	100.0	787	10	US-09-984-186-16
42	3103	100.0	787	15	US-10-237-667-16
43	3103	100.0	787	15	US-10-237-708-16
44	3103	100.0	787	15	US-10-237-866-16
45	3103	100.0	787	15	US-10-237-866-16

ALIGNMENTS

RESULT :

```
US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20020123080A1
; GENERAL INFORMATION:
; APPLICANT: Sommerschein, Carlos
; Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DCS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/929,552
; FILING DATE: 14-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/769,746
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
```

RESULT 2
 S-09-932-613-445
 Sequence 445, Application US/09932613
 Publication No. US20030091565A1
 GENERAL INFORMATION:
 APPLICANT: Human Genome Sciences, Inc.
 APPLICANT: Beltzer, James P.
 APPLICANT: Potter, M. Daniel
 APPLICANT: Fleming, Tony J.
 APPLICANT: Rosen, Craig A.
 TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
 FILE REFERENCE: DXX-025.1 PCT: DXX-025.1 US
 CURRENT APPLICATION NUMBER: US/09/932,613
 CURRENT FILING DATE: 2001-08-17
 NUMBER OF SEQ ID NOS: 458
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 445
 LENGTH: 585
 TYPE: PRT
 ORGANISM: HomoSapiens
 S-09-932-613-445

; FILING DATE: 19-DEC-1996
 ; INFORMATION FOR SEQ ID NO: 26:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 585 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: <Unknown>
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
 US-09-984-010-26

Query Match 100.0%; Score 3103; DB 11; Length 585;
 Best Local Similarity 100.0%; Pred. No. 2,4e-269;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHFKDLGEENFKALVLIAPAQYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHSEVAHFKDLGEENFKALVLIAPAQYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFQHKDDNPRLVLRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFQHKDDNPRLVLRPEV 120
 QY 121 DMCTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKYKAAFTCCQAADKAACLLP 180
 DB 121 DMCTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKYKAAFTCCQAADKAACLLP 180
 QY 181 KLDELREGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
 DB 181 KLDELREGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
 DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYVARRHPDYSVLLLRJAKTYETTLK 360
 DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYVARRHPDYSVLLLRJAKTYETTLK 360
 QY 361 CAADPHCEYAKVFEFKPLVEEPQNLIKNCELFEQLGEYKFNALLVRYTKKVPQVST 420
 DB 361 CAADPHCEYAKVFEFKPLVEEPQNLIKNCELFEQLGEYKFNALLVRYTKKVPQVST 420
 QY 421 PTLVEVSRLNGVSKCKCKHPEAKRMPCAEEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRLNGVSKCKCKHPEAKRMPCAEEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDVETVYPKEFNAETTFHADICTLSEKQIKKOTALVELVKKHKPKAT 540
 DB 481 LVNRRPCFSALEVDVETVYPKEFNAETTFHADICTLSEKQIKKOTALVELVKKHKPKAT 540
 QY 541 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEEGKKLVAASQAALGJ 585
 DB 541 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEEGKKLVAASQAALGJ 585

RESULT 4
 US-09-833-041-18
 ; Sequence 18, Application US/09833041
 ; Publication No. US20030125247A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen, Craig A.
 ; APPLICANT: Haseltine, William A.
 ; TITLE OF INVENTION: Albumin Fusion Proteins
 ; FILE REFERENCE: PF545
 ; CURRENT APPLICATION NUMBER: US/09/833,041
 ; CURRENT FILING DATE: 2001-04-12
 ; PRIOR APPLICATION NUMBER: 60/229,358
 ; PRIOR FILING DATE: 2000-04-12
 ; PRIOR APPLICATION NUMBER: 60/256,931
 ; PRIOR FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/199,384
 ; PRIOR FILING DATE: 2000-04-25
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 18
 ; LENGTH: 585
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-833-041-18
 Query Match 100.0%; Score 3103; DB 11; Length 585;
 Best Local Similarity 100.0%; Pred. No. 2,4e-269;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHFKDLGEENFKALVLIAPAQYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHSEVAHFKDLGEENFKALVLIAPAQYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFQHKDDNPRLVLRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFQHKDDNPRLVLRPEV 120
 QY 121 DMCTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKYKAAFTCCQAADKAACLLP 180
 DB 121 DMCTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKYKAAFTCCQAADKAACLLP 180
 QY 181 KLDELREGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
 DB 181 KLDELREGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
 DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYVARRHPDYSVLLLRJAKTYETTLK 360
 DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYVARRHPDYSVLLLRJAKTYETTLK 360
 QY 361 CAADPHCEYAKVFEFKPLVEEPQNLIKNCELFEQLGEYKFNALLVRYTKKVPQVST 420
 DB 361 CAADPHCEYAKVFEFKPLVEEPQNLIKNCELFEQLGEYKFNALLVRYTKKVPQVST 420
 QY 421 PTLVEVSRLNGVSKCKCKHPEAKRMPCAEEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRLNGVSKCKCKHPEAKRMPCAEEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDVETVYPKEFNAETTFHADICTLSEKQIKKOTALVELVKKHKPKAT 540
 DB 481 LVNRRPCFSALEVDVETVYPKEFNAETTFHADICTLSEKQIKKOTALVELVKKHKPKAT 540
 QY 541 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEEGKKLVAASQAALGJ 585
 DB 541 KEQLKAVMDDFAAFVEKCKCCKADDDKTCFAEEGKKLVAASQAALGJ 585

RESULT 5
 US-10-153-604A-5
 ; Sequence 5, Application US/10:53604A
 ; Publication No. US20030:43191A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bell et al.
 ; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
 ; FILE REFERENCE: PF556
 ; CURRENT APPLICATION NUMBER: US/10/153,604A
 ; CURRENT FILING DATE: 2002-05-24
 ; PRIOR APPLICATION NUMBER: 60/293,212
 ; PRIOR FILING DATE: 2001-05-25
 ; NUMBER OF SEQ ID NOS: 137
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 585
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens

US-10-153-604A-5

```
Query Match      100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRFKDLGEENFKALVLIAPAYLQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRRFKDLGEENFKALVLIAPAYLQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDPNLPRLVREVEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDPNLPRLVREVEV 120
QY 121 DVMTCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 180
DB 121 DVMTCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKASLQKFGGERAFKAWAVARLSORFPKAEFAEVSUKLVTDLTG 240
DB 181 KLDELDEGKASSAKORLKASLQKFGGERAFKAWAVARLSORFPKAEFAEVSUKLVTDLTG 240
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAAKQVFLGMFLVEYARRHPDYVSVLLLRLLAKTYETTLEK 360
DB 301 DLPSLAADFVESKDVCKNYAAKQVFLGMFLVEYARRHPDYVSVLLLRLLAKTYETTLEK 360
QY 361 CAADPHCEYAKVDFEPLVEEPONLIKONCELFGEOLGEYKFNALLVRYTKKVPQVST 420
DB 361 CAADPHCEYAKVDFEPLVEEPONLIKONCELFGEOLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGVGSKCKCHPEAKMPCAEYDLSVNLQCLVLEHKTVPVSDRVTKCTTES 480
DB 421 PTLVEVSRNLGVGSKCKCHPEAKMPCAEYDLSVNLQCLVLEHKTVPVSDRVTKCTTES 480
QY 481 LVNRRPCFSALVEVDITYVPKEFNAETFTFHADICTLSEKERQIKKQALVELVVKHKPKAT 540
DB 481 LVNRRPCFSALVEVDITYVPKEFNAETFTFHADICTLSEKERQIKKQALVELVVKHKPKAT 540
QY 541 KEQLKAVMDDFAAFEVKCKKADDDKTCFAEEGKKLVAAASQAALGL 585
DB 541 KEQLKAVMDDFAAFEVKCKKADDDKTCFAEEGKKLVAAASQAALGL 585
```

RESULT 6

```
US-09-833-117-18
; Sequence 18, Application US/09833117
; Publication No. US20030171267A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF543
; CURRENT APPLICATION NUMBER: US/09/833,117
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18
```

Query Match

100.0%; Score 3103; DB 12; Length 585;

```
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 DAHSEVAHRRFKDLGEENFKALVLIAPAYLQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRRFKDLGEENFKALVLIAPAYLQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDPNLPRLVREVEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDPNLPRLVREVEV 120
QY 121 DVMTCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 180
DB 121 DVMTCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKASLQKFGGERAFKAWAVARLSORFPKAEFAEVSUKLVTDLTG 240
DB 181 KLDELDEGKASSAKORLKASLQKFGGERAFKAWAVARLSORFPKAEFAEVSUKLVTDLTG 240
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAAKQVFLGMFLVEYARRHPDYVSVLLLRLLAKTYETTLEK 360
DB 301 DLPSLAADFVESKDVCKNYAAKQVFLGMFLVEYARRHPDYVSVLLLRLLAKTYETTLEK 360
QY 361 CAADPHCEYAKVDFEPLVEEPONLIKONCELFGEOLGEYKFNALLVRYTKKVPQVST 420
DB 361 CAADPHCEYAKVDFEPLVEEPONLIKONCELFGEOLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGVGSKCKCHPEAKMPCAEYDLSVNLQCLVLEHKTVPVSDRVTKCTTES 480
DB 421 PTLVEVSRNLGVGSKCKCHPEAKMPCAEYDLSVNLQCLVLEHKTVPVSDRVTKCTTES 480
QY 481 LVNRRPCFSALVEVDITYVPKEFNAETFTFHADICTLSEKERQIKKQALVELVVKHKPKAT 540
DB 481 LVNRRPCFSALVEVDITYVPKEFNAETFTFHADICTLSEKERQIKKQALVELVVKHKPKAT 540
QY 541 KEQLKAVMDDFAAFEVKCKKADDDKTCFAEEGKKLVAAASQAALGL 585
DB 541 KEQLKAVMDDFAAFEVKCKKADDDKTCFAEEGKKLVAAASQAALGL 585
```

RESULT 7

```
US-10-319-263-1
; Sequence 1, Application US/10319263
; Publication No. US20030180623A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISC007
; CURRENT APPLICATION NUMBER: US/10/319,263
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-10-319-263-1

```
Query Match      100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRRFKDLGEENFKALVLIATFAQYQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 2 DAHKEVAHRRFKDLGEENFKALVLIATFAQYQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDKNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDKNPNLRLVRPEV 120
QY 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAKAAACLLP 180
DB 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAKAAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDRDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSTSSKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECADRADLAKYICENQDSTSSKLKECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADPVESKDVCKNYAEAKQVFLGMFLYIYARRHPDYVSWLLRLAKTYETTLK 360
DB 301 DLPSLAADPVESKDVCKNYAEAKQVFLGMFLYIYARRHPDYVSWLLRLAKTYETTLK 360
QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKQNCLELFEQGEYKFNALLVRYTKVPOVST 420
DB 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKQNCLELFEQGEYKFNALLVRYTKVPOVST 420
QY 421 PTLVEVSRNLGVKSGKCKKHPEAKMPCAEYLSVWLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGVKSGKCKKHPEAKMPCAEYLSVWLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKEROIKKQTAALVELVKHKPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKEROIKKQTAALVELVKHKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAAASQAALGL 585
DB 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAAASQAALGL 585
```

RESULT 8

```
US-10-319-263-2
; Sequence 2, Application US/103:9263
; Publication No. US20030180820A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: IS0007
; CURRENT APPLICATION NUMBER: US/10/319,263
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 585
; TYPE: PRT
```

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: MOD_RES

LOCATION: (1)..(585)

OTHER INFORMATION: ACETYLATION

US-10-319-263-2

```
Query Match      100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRRFKDLGEENFKALVLIATFAQYQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 2 DAHKEVAHRRFKDLGEENFKALVLIATFAQYQQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDKNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDKNPNLRLVRPEV 120
QY 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAKAAACLLP 180
DB 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAKAAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDRDEGKASSAKQRLKCAASLOKFGGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSTSSKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECADRADLAKYICENQDSTSSKLKECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADPVESKDVCKNYAEAKQVFLGMFLYIYARRHPDYVSWLLRLAKTYETTLK 360
DB 301 DLPSLAADPVESKDVCKNYAEAKQVFLGMFLYIYARRHPDYVSWLLRLAKTYETTLK 360
QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKQNCLELFEQGEYKFNALLVRYTKVPOVST 420
DB 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKQNCLELFEQGEYKFNALLVRYTKVPOVST 420
QY 421 PTLVEVSRNLGVKSGKCKKHPEAKMPCAEYLSVWLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGVKSGKCKKHPEAKMPCAEYLSVWLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKEROIKKQTAALVELVKHKPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKEROIKKQTAALVELVKHKPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAAASQAALGL 585
DB 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAAASQAALGL 585
```

RESULT 9

```
US-10-414-469-1
; Sequence 1, Application US/104:4469
; Publication No. US2003019063A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: IS0007
; CURRENT APPLICATION NUMBER: US/10/414,469
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 09/806,247
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: PCT/US99/22905
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
```

; PRIOR APPLICATION NUMBER: 09/165,926
 ; PRIOR FILING DATE: 1998-10-02
 ; PRIOR APPLICATION NUMBER: 09/165,581
 ; PRIOR FILING DATE: 1998-10-02
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 585
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-414-469-1

Query Match 100.0%; Score 3103; DB 12; Length 585;
 Best Local Similarity 100.0%; Pred. No. 2.4e-269;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 3;

QY	1	DAHKSEVAHRFKDGLGDNFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAS	60
DB	1	DAHKSEVAHRFKDGLGDNFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAS	60
QY	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVPRPV	120
DB	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVPRPV	120
QY	121	DVMCTAFPHNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP	180
DB	121	DVMCTAFPHNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP	180
QY	181	KLDELDEGKASSAKQRLKASLQKGFGERAFKAWAVARLSORFFKAEFAEVSKLVTDLTK	240
DB	181	KLDELDEGKASSAKQRLKASLQKGFGERAFKAWAVARLSORFFKAEFAEVSKLVTDLTK	240
QY	241	VHTECHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA	300
DB	241	VHTECHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA	300
QY	301	DLPSLAADFVSEKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLLRKAKTYETTLK	360
DB	301	DLPSLAADFVSEKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLLRKAKTYETTLK	360
QY	361	CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQGLGEYKFNALLVRYTKKVPQVST	420
DB	361	CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQGLGEYKFNALLVRYTKKVPQVST	420
QY	421	PTLVEVSRLNGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKKCTES	480
DB	421	PTLVEVSRLNGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKKCTES	480
QY	481	LNNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKHKPKAT	540
DB	481	LNNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKHKPKAT	540
QY	541	KEQLKAVMDDFAAAFVEKCKCKADDKETCFABEGKKLVAAASQAALGL	585
DB	541	KEQLKAVMDDFAAAFVEKCKCKADDKETCFABEGKKLVAAASQAALGL	585

RESULT 10
 US-10-414-469-2
 ; Sequence 2, Application US/104:4469
 ; Publication No. US20030190691A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bar-Or M.D., David
 ; APPLICANT: Lau Ph.D., Edward
 ; APPLICANT: Winkler M.D., James V.
 ; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
 ; TITLE OF INVENTION: Kits
 ; FILE REFERENCE: ISCC007
 ; CURRENT APPLICATION NUMBER: US/10/414,469
 ; CURRENT FILING DATE: 2003-04-15
 ; PRIOR APPLICATION NUMBER: 09/806,247
 ; PRIOR FILING DATE: 1999-10-01
 ; PRIOR APPLICATION NUMBER: PCT/US99/22905

; PRIOR FILING DATE: 1999-10-01
 ; PRIOR APPLICATION NUMBER: 60/115,392
 ; PRIOR FILING DATE: 1999-01-11
 ; PRIOR APPLICATION NUMBER: 60/102,738
 ; PRIOR FILING DATE: 1998-10-02
 ; PRIOR APPLICATION NUMBER: 09/165,926
 ; PRIOR FILING DATE: 1998-10-02
 ; PRIOR APPLICATION NUMBER: 09/165,581
 ; PRIOR FILING DATE: 1998-10-02
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 585
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURES:
 ; NAME/KEY: MOD RES
 ; LOCATION: (1)..(585)
 ; OTHER INFORMATION: ACETYLATION
 ; US-10-414-469-2

Query Match 100.0%; Score 3103; DB 12; Length 585;
 Best Local Similarity 100.0%; Pred. No. 2.4e-269;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DAHKSEVAHRFKDGLGDNFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAS	60
DB	1	DAHKSEVAHRFKDGLGDNFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAS	60
QY	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVPRPV	120
DB	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVPRPV	120
QY	121	DVMCTAFPHNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP	180
DB	121	DVMCTAFPHNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAADKAAACLLP	180
QY	181	KLDELDEGKASSAKQRLKASLQKGFGERAFKAWAVARLSORFFKAEFAEVSKLVTDLTK	240
DB	181	KLDELDEGKASSAKQRLKASLQKGFGERAFKAWAVARLSORFFKAEFAEVSKLVTDLTK	240
QY	241	VHTECHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA	300
DB	241	VHTECHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA	300
QY	301	DLPSLAADFVSEKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLLRKAKTYETTLK	360
DB	301	DLPSLAADFVSEKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLLRKAKTYETTLK	360
QY	361	CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQGLGEYKFNALLVRYTKKVPQVST	420
DB	361	CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQGLGEYKFNALLVRYTKKVPQVST	420
QY	421	PTLVEVSRLNGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKKCTES	480
DB	421	PTLVEVSRLNGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKKCTES	480
QY	481	LNNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKHKPKAT	540
DB	481	LNNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKKHKPKAT	540
QY	541	KEQLKAVMDDFAAAFVEKCKCKADDKETCFABEGKKLVAAASQAALGL	585
DB	541	KEQLKAVMDDFAAAFVEKCKCKADDKETCFABEGKKLVAAASQAALGL	585

RESULT 11
 US-09-932-322-445
 ; Sequence 445, Application US/09932322
 ; Publication No. US20030194743A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dyac Corp.
 ; APPLICANT: Beltzer, James P.

APPLICANT: Potter, M. Daniel;
APPLICANT: Fleming, Tony J.
APPLICANT: Ladner, Robert Charles
TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (B-lys)
FILE REFERENCE: DYA-018.1 PCT; DYA-018.1 US
CURRENT APPLICATION NUMBER: US/09/932,322
CURRENT FILING DATE: 2001-08-17
NUMBER OF SEQ ID NOS: 458
SOFTWARE: PatentIn version 3.1
SEQ ID NO 445
LENGTH: 585
TYPE: PRT
ORGANISM: HomoSapiens
US-09-932-322-445

Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60
DB 1 DAHKEVAHRFKDGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPRLVREPV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPRLVREPV 120
QY 121 DVMCTAFHNDNEETFLKKYLYEIARHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 121 DVMCTAFHNDNEETFLKKYLYEIARHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELDEGKASSAKORLKASLOKFGGERAFKAMAVARLSORFFKAEPAEVSUKLVTDLT 240
DB 181 KLDELDEGKASSAKORLKASLOKFGGERAFKAMAVARLSORFFKAEPAEVSUKLVTDLT 240
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKJKECCCKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTTECHGDLLECADRADLAKYICENQDSISSKJKECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSSAAADFVESKDVCKNYAEAKOVFLGMFLYEAARRHPDYSVLLLRKATYETTLK 360
DB 301 DLPSSAAADFVESKDVCKNYAEAKOVFLGMFLYEAARRHPDYSVLLLRKATYETTLK 360
QY 361 CAADAPHECYAKVDFEFPKLVSEPNLIKQNCLEFEQGLGEYKFNQALVRYTKKVPQVST 420
DB 361 CAADAPHECYAKVDFEFPKLVSEPNLIKQNCLEFEQGLGEYKFNQALVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEEDYLSVNLQCLCVLHEKTPVSDRVTKCTES 480
DB 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEEDYLSVNLQCLCVLHEKTPVSDRVTKCTES 480
QY 481 LVNRRPCFSALEVDYTPVKFENAEFTTFHADICTLSEKERQIKKQALVELVGHKPKAT 540
DB 481 LVNRRPCFSALEVDYTPVKFENAEFTTFHADICTLSEKERQIKKQALVELVGHKPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGSKLVAAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGSKLVAAASQAALGL 585

RESULT 12
US-10-413-831-1
Sequence 1, Application US/10413831
Publication No. US20030194813A1
GENERAL INFORMATION:
APPLICANT: Bar-Or M.D., David
APPLICANT: Lau Ph.D., Edward
APPLICANT: Winkler M.D., James V.
TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and Kits
FILE REFERENCE: ISC007
CURRENT APPLICATION NUMBER: US/10/413,831
CURRENT FILING DATE: 2003-04-15

PRIOR APPLICATION NUMBER: US/09/806,247
PRIOR FILING DATE: 2001-07-16
PRIOR APPLICATION NUMBER: 60/115,392
PRIOR FILING DATE: 1999-01-11
PRIOR APPLICATION NUMBER: 60/102,738
PRIOR FILING DATE: 1998-10-02
PRIOR APPLICATION NUMBER: 09/165,926
PRIOR FILING DATE: 1998-10-02
PRIOR APPLICATION NUMBER: 09/165,581
PRIOR FILING DATE: 1998-10-02
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 585
LENGTH: 585
TYPE: PRT
ORGANISM: Homo sapiens
US-10-413-831-1

Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60
DB 1 DAHKEVAHRFKDGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPRLVREPV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPRLVREPV 120
QY 121 DVMCTAFHNDNEETFLKKYLYEIARHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 121 DVMCTAFHNDNEETFLKKYLYEIARHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELDEGKASSAKORLKASLOKFGGERAFKAMAVARLSORFFKAEPAEVSUKLVTDLT 240
DB 181 KLDELDEGKASSAKORLKASLOKFGGERAFKAMAVARLSORFFKAEPAEVSUKLVTDLT 240
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKJKECCCKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTTECHGDLLECADRADLAKYICENQDSISSKJKECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSSAAADFVESKDVCKNYAEAKOVFLGMFLYEAARRHPDYSVLLLRKATYETTLK 360
DB 301 DLPSSAAADFVESKDVCKNYAEAKOVFLGMFLYEAARRHPDYSVLLLRKATYETTLK 360
QY 361 CAADAPHECYAKVDFEFPKLVSEPNLIKQNCLEFEQGLGEYKFNQALVRYTKKVPQVST 420
DB 361 CAADAPHECYAKVDFEFPKLVSEPNLIKQNCLEFEQGLGEYKFNQALVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEEDYLSVNLQCLCVLHEKTPVSDRVTKCTES 480
DB 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEEDYLSVNLQCLCVLHEKTPVSDRVTKCTES 480
QY 481 LVNRRPCFSALEVDYTPVKFENAEFTTFHADICTLSEKERQIKKQALVELVGHKPKAT 540
DB 481 LVNRRPCFSALEVDYTPVKFENAEFTTFHADICTLSEKERQIKKQALVELVGHKPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGSKLVAAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGSKLVAAASQAALGL 585

RESULT 13
US-10-413-831-2
Sequence 2, Application US/10413831
Publication No. US20030194813A1
GENERAL INFORMATION:
APPLICANT: Bar-Or M.D., David
APPLICANT: Lau Ph.D., Edward
APPLICANT: Winkler M.D., James V.
TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and Kits

```

; FILE REFERENCE: ISCO07
; CURRENT APPLICATION NUMBER: US/10/413,831
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: US/09/806,247
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,739
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (1)-(585)
; OTHER INFORMATION: ACETYLATION
US-10-413-831-2

```

```

Query Match: 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHREFKDLGGENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHREFKDLGGENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVREPV 120
DB 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVREPV 120
QY 121 DVMTAFHNDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACL 180
DB 121 DVMTAFHNDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACL 180
QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
DB 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
QY 241 VHTCCGHDLLCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDKMPA 300
DB 241 VHTCCGHDLLCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDKMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLGYEYARRHPDYSVVLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLGYEYARRHPDYSVVLRLAKTYETTLK 360
QY 361 CAADPHCEYAKVDFDEPKPLVEEPQNLKONCELFQGLGEYKFNALVRYTKVPQVST 420
DB 361 CAADPHCEYAKVDFDEPKPLVEEPQNLKONCELFQGLGEYKFNALVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYVVPKFEFNAETTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540
DB 481 LVNRRPCFSALEVDYVVPKFEFNAETTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540
QY 541 KEQLKAVMDFFAAFEVKCCCKADDDKCTCFABEGKLVAAASQAALGL 585
DB 541 KEQLKAVMDFFAAFEVKCCCKADDDKCTCFABEGKLVAAASQAALGL 585

```

```

RESULT 14
US-10-153-064-5
; Sequence 5, Application US/10/53064

```

```

; Publication No. US200201428-4A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-153-064-5

Query Match: 100.0%; Score 3103; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHREFKDLGGENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHREFKDLGGENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVREPV 120
DB 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVREPV 120
QY 121 DVMTAFHNDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACL 180
DB 121 DVMTAFHNDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACL 180
QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
DB 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
QY 241 VHTCCGHDLLCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDKMPA 300
DB 241 VHTCCGHDLLCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDKMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLGYEYARRHPDYSVVLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLGYEYARRHPDYSVVLRLAKTYETTLK 360
QY 361 CAADPHCEYAKVDFDEPKPLVEEPQNLKONCELFQGLGEYKFNALVRYTKVPQVST 420
DB 361 CAADPHCEYAKVDFDEPKPLVEEPQNLKONCELFQGLGEYKFNALVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYVVPKFEFNAETTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540
DB 481 LVNRRPCFSALEVDYVVPKFEFNAETTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540
QY 541 KEQLKAVMDFFAAFEVKCCCKADDDKCTCFABEGKLVAAASQAALGL 585
DB 541 KEQLKAVMDFFAAFEVKCCCKADDDKCTCFABEGKLVAAASQAALGL 585

RESULT 15
US-09-984-010-7
; Sequence 7, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
; AND SERUM ALBUMIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & JUNNER, LLP

```

STREET: 1300 I Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,010
FILING DATE: 21-May-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/091,873
FILING DATE: 25-JUN-1998
APPLICATION NUMBER: PCT/GB96/03164
FILING DATE: 19-DEC-1996
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 604 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-984-010-7
Query Match 100.0%; Score 3103; DB 1; Length 604;
Best Local Similarity 100.0%; Pred. No. 2,5e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRPKDGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
DB 20 DAHKSEVAHRPKDGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 79
QY 61 NCKSLHTLFGDKCTVATLRETYGEMADCCAKOEPERNECFQHKDDNPNLRLVRPEV 120
DB 80 NCKSLHTLFGDKCTVATLRETYGEMADCCAKOEPERNECFQHKDDNPNLRLVRPEV 139
QY 121 DVNCTAFHDNEETFLKYLVEIARRHPYFYAPELFFAKRYKAFTCCQAADKAACLLP 180
DB 140 DVNCTAFHDNEETFLKYLVEIARRHPYFYAPELFFAKRYKAFTCCQAADKAACLLP 199
QY 191 KLDLRDEGKASSAKQRLKASLOKQGERAFKAWAVARLSQRPFKARFAYSKLVTDJTK 240
DB 200 KLDLRDEGKASSAKQRLKASLOKQGERAFKAWAVARLSQRPFKARFAYSKLVTDJTK 259
QY 241 VHTCCCHGDILLECADDRADLAKYICENQDSISSKLCCKEKPJLEKSHCAEVENDEMPA 300
DB 260 VHTCCCHGDILLECADDRADLAKYICENQDSISSKLCCKEKPJLEKSHCAEVENDEMPA 319
QY 301 DLPSLAADFVESKQCKNYAEAKQVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLEKC 360
DB 320 DLPSLAADFVESKQCKNYAEAKQVFLGMFLYEVARRHPDYSVVLLRLAKTYETTLEKC 379
QY 361 CAADPHCEYAKVFEDEKPLVEEZEQNLIKQNCELFEQLGEYKQNALIVRYTKKVPQVST 420
DB 380 CAADPHCEYAKVFEDEKPLVEEZEQNLIKQNCELFEQLGEYKQNALIVRYTKKVPQVST 439
QY 421 PTLVEVSRNLKGVSKCKHPFAKMPCAEDYLSVNLNCLCVLHEKTPVSDRVTKCTTES 480
DB 440 PTLVEVSRNLKGVSKCKHPFAKMPCAEDYLSVNLNCLCVLHEKTPVSDRVTKCTTES 499
QY 481 LVNRRPCFSALVEDETVVPKEFNAETFTFHADICTLSEKSRQKKQTALVELVKKHFKAT 540
DB 500 LVNRRPCFSALVEDETVVPKEFNAETFTFHADICTLSEKSRQKKQTALVELVKKHFKAT 559
QY 541 KEQLKAVNDDPFAAFVEKCKCKADCKETCFABEGKKLVAASQAALGJ 585
DB 560 KEQLKAVNDDPFAAFVEKCKCKADCKETCFABEGKKLVAASQAALGJ 604

Search completed: October 27, 2003, 15:31:04
Job time : 74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 15:15:53 : Search time 43 seconds
(without alignments)
1308.341 Million cell updates/sec

Title: US-09-833-117-18

Perfect score: 3103

Sequence: 1 DAHKSEVAHFKDLGEENFK.....TCFAEECKKLVAASQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168692 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 3

Maximum DB seq length: 2300000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_76.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	3103	100.0	609	1	ABHUS		serum albumin prec
2	2942	94.8	600	2	A47391		serum albumin prec
3	2620	84.4	608	2	S57632		serum albumin prec
4	2475.5	79.8	607	1	ABHOS		serum albumin prec
5	2446.5	78.8	607	1	ABHOS		serum albumin prec
6	2432.5	78.4	607	1	ABHOS		serum albumin prec
7	2426	78.2	608	1	ABRTS		serum albumin prec
8	2411.5	77.7	605	1	ABPGS		serum albumin prec
9	2387	76.9	609	2	CS5838		albumin - Xongolia
10	1861	60.3	453	2	A51339		serum albumin - mo
11	1557.5	50.2	615	1	ASCHS		serum albumin prec
12	1253.5	40.4	609	2	UC4258		alpha-fetoprotein
13	1249.5	40.3	609	1	PFHU		alpha-fetoprotein
14	1242.5	40.0	609	1	PFHO		alpha-fetoprotein
15	1205	38.8	607	1	ASXJ72		74K albumin prec
16	1181.5	38.1	265	2	I46986		albumin - dog liver
17	1175.5	37.9	608	2	ASXJ68		68K serum albumin
18	1084	34.9	605	1	FPMS		alpha-fetoprotein
19	1067	34.4	611	1	FPRT		alpha-fetoprotein
20	1055	34.0	599	1	A54906		afamin precursor
21	928.5	29.9	614	2	S59527		serum albumin prec
22	928	29.9	608	2	A53195		afamin precursor
23	747.5	24.1	608	1	ABONS1		serum albumin 1 pr
24	742.5	23.9	608	1	ABONS2		serum albumin 2 pr
25	699	22.5	382	2	A37253		serum albumin - bu
26	440.5	14.2	423	1	S27941		serum albumin - se
27	386	12.4	474	1	VYHVD		vitamin D-binding
28	385	12.4	476	1	VYR7D		vitamin D-binding
29	372	12.0	472	1	A35327		vitamin D-binding

ALIGNMENTS

RESULT 1

ABHUS

serum albumin precursor [validated] - human

N;Alternate names: preproalbumin

N;Contains: kinetensin

C;Species: Homo sapiens (man)

C;Date: 29-Jul-1981 #sequence, revision 31-Jan-1997 #text change 17-Mar-2000

C;Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S56422;

R;Jawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; See

Nucleic Acids Res. 9, 6103-6114, 1981

A;Title: The sequence of human serum albumin cDNA and its expression in Escherichia

A;Reference number: A93743; MUID:82081892; PMID:6171778

A;Accession: A93743

A;Molecule type: mRNA

A;Residues: 1-419, 'K', 421-609 <LAW>

A;Cross-references: EMBL:V00495; GB:J00078; GB:J00133; GB:J00133; NID:G28591; PIDN:CF

R;Dugaiczky, A.; Law, S.W.; Dennison, O.E.

Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982

A;Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.

A;Reference number: A93936; MUID:82105994; PMID:6275391

A;Accession: A93936

A;Molecule type: mRNA

A;Residues: 1-120, 'G', 122-609 <DUG>

A;Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; P;D:G28590

R;Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.

J. Biol. Chem. 261, 3244-3251, 1986

A;Title: The human albumin gene. Characterization of the 5' and 3' flanking regions

A;Reference number: I39427; MUID:86140099; PMID:2419329

A;Accession: I39427

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-26 <URA>

A;Cross-references: GB:J13075; NID:G178330; PIDN:AAAS1688.1; PID:G553173

R;Watkins, S.; Radison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.

Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994

A;Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian fami

A;Reference number: I59286; MUID:94181575; PMID:8134387

A;Accession: I59286

A;Status: translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 282-290, 'KSRFDLC', <WAT>

A;Cross-references: GB:S69192; NID:G546032; PIDN:AAB30282.1; PID:G546033

A;Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia

R;Radison, S.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putra

Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994

A;Title: Genetic variants of human serum albumin in Italy: point mutants and a carbox

A;Reference number: I59313; MUID:94294404; PMID:8022807

A;Accession: I59313

A;Status: translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 589-590, 'ALPRRVKMLLOVKLP', <MAL>

A;Cross-references: GB:S70799; NID:G547231; PIDN:AAB31177.1; P;D:G547232

cag island protein
cag pathogenicity
calcium-binding pr
hypothetical prote
major surface glyco
cell surface glyco
hypothetical prote
sperm tail-specific
kinesin homolog F2
major surface glyco
myosin heavy chain
cell-cycle-depende
glycoprotein A - p
hypothetical prote
giantin - human
embryonic muscle m

A:Note: this frame-shift variant is designated albumin Bazzano; four additional variants
 R:Menaya, J.; Parrilla, R.; Ayuso, M.S.
 submitted to the EMBL Data Library, March 1995
 A:Reference number: G08292
 A:Accession: G01747
 A>Status: translated from GB/EMBL/DBSJ
 A:Molecule type: mRNA
 A:Residues: 1-120, 'G', 122-455 <MEN>
 A:Cross-references: EMBL:U22961; NID:G763428; PID:AAA64922.1; PID:G763431
 R:Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
 Biochem. J. 308, 321-325, 1995
 A:Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2
 A:Reference number: S55314; MUID:95275251; PMID:7755581
 A:Accession: S55314
 A:Molecule type: protein
 A:Residues: 19-27 <LED>
 R:Meioun, B.; Moravsek, L.; Kostka, V.
 FEBS Lett. 58, 134-137, 1975
 A:Title: Complete amino acid sequence of human serum albumin.
 A:Reference number: A9420; MUID:76187907; PMID:1225573
 A:Accession: A9420
 A:Molecule type: protein
 A:Residues: 25-117, EQ, 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-
 R:Roehr, U.; Spitteler, G.; Tripler, E.
 Justus Liebig's Ann. Chem. 9, 881-884, 1988
 A:Title: Isolation and structure elucidation of middle-molecular weight peptides from u
 A:Reference number: S06422
 A:Note: this paper is in German, with an English abstract
 A:Accession: S06422
 A:Molecule type: protein
 A:Residues: 25-48 <ROE>
 R:Pinch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
 Arch. Biochem. Biophys. 305, 595-599, 1993
 A:Title: Mass spectrometric identification of modifications to human serum albumin treat
 A:Reference number: S38882; MUID:93394321; PMID:8373198
 A:Accession: S38882
 A:Molecule type: protein
 A:Residues: 45-67; 141-160; 311-337; 469-490; 570-581 <PIN>
 R:Kausler, E.; Spitteler, G.
 Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
 A:Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mitreimol
 A:Reference number: S17599; MUID:92126241; PMID:1772598
 A:Accession: S17599
 A:Molecule type: protein
 A:Residues: 25-54; 354-357; 431-447 <KAJ>
 A:Note: 49-Leu was also found
 R:Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
 J. Immunol. 143, 1680-1684, 1989
 A:Title: Structures of histamine-releasing peptides formed by the action of acid proteas
 A:Reference number: A45800; MUID:89341436; PMID:2474609
 A:Accession: A45800
 A:Molecule type: protein
 A:Residues: 166-173 <CAR>
 R:Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa
 Biochem. Biophys. Res. Commun. 136, 983-988, 1986
 A:Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tre
 A:Reference number: A03239; MUID:86242180; PMID:3087352
 A:Accession: A03239
 A:Molecule type: protein
 A:Residues: 166-173, 'L' <XOG>
 R:Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S
 Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
 A:Title: Mutations in genetic variants of human serum albumin found in Italy.
 A:Reference number: A38255; MUID:91062352; PMID:2247440
 A:Accession: C38255
 A:Molecule type: protein
 A:Residues: 76-111 <GAL1>
 A:Accession: B38255
 A:Molecule type: protein
 A:Residues: 82-105, 'K', 107-110 <GAL2>
 A:Note: this variant is designated albumin Vibo Valentia
 A:Accession: A38255
 A:Molecule type: protein

A:Residues: 76-93, 'K', 85-156 <GAL3>
 A:Note: this variant is designated albumin Torino
 R:Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
 Eur. J. Biochem. 214, 437-444, 1993
 A:Title: The structural characterization and bilirubin-binding properties of albumin
 A:Reference number: S33298; MUID:93292504; PMID:8513793
 A:Accession: S33298
 A:Molecule type: protein
 A:Residues: 255-263, 'E', 265-281 <MIN1>
 A:Note: this variant is designated albumin Herborn
 R:Minchiotti, L.; Galliano, M.; Scoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Po
 Biochim. Biophys. Acta 1119, 232-238, 1992
 A:Title: Two albumins with identical electrophoretic mobility are produced by dif
 A:Reference number: S21078; MUID:92190239; PMID:1347703
 A:Accession: S21078
 A:Molecule type: protein
 A:Residues: 354-356, 'K', 358-378 <MIN2>
 A:Note: this variant is designated albumin Sondrio; another variant Paris-2 is repor
 R:He, X.M.; Carter, D.C.
 Nature 358, 209-215, 1992
 A:Title: Atomic structure and chemistry of human serum albumin.
 A:Reference number: A46756; MUID:92334427; PMID:1630489
 A:Contents: annotation; x-ray crystallography, 2.8 angstroms
 R:Brown, J.R.; Stockley, P.; Behrens, P.O.
 In The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23
 A:Reference number: A94442
 A:Contents: annotation; three-dimensional structure and disulfide bonds
 R:Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meioun, B.
 Collect. Czech. Chem. Commun. 42, 564-579, 1977
 A:Title: Disulfide bonds in human serum albumin.
 A:Reference number: A90930
 A:Contents: annotation; disulfide bonds
 R:Jacobsen, C.
 Biochem. J. 17, 453-459, 1978
 A:Title: Lysine residue 240 of human serum albumin is involved in high-affinity bind
 A:Reference number: A90299; MUID:78186630; PMID:656055
 A:Contents: annotation; bilirubin-binding site
 R:Peters, T.; Reed, R.G.
 In Albumin: Structure, Biosynthesis, Function, Peters, J., and Scholm, I., eds., 11
 A:Title: Serum albumin: conformation and active sites.
 A:Reference number: A94408
 A:Contents: annotation; binding sites
 R:Harper, M.E.; Dugaczky, A.
 Am. J. Hum. Genet. 35, 565-572, 1983
 A:Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein g
 A:Reference number: A90028; MUID:8329982; PMID:6192711
 A:Contents: annotation; gene position
 R:Walker, J.E.
 FEBS Lett. 66, 173-175, 1976
 A:Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic a
 A:Reference number: A46755; MUID:76257808; PMID:955075
 A:Contents: annotation
 A:Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic
 R:Bonney, J.P.; Fonda, M.L.; Feldhoff, R.C.
 FEBS Lett. 298, 266-268, 1992
 A:Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-pho
 A:Reference number: A56294; MUID:92183881; PMID:1544460
 A:Contents: annotation
 A:Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described;
 A:ase activity
 C:Comment: Serum albumin, a predominant protein in the plasma of adults, is synthe
 Lirubin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (w
 C:Comment: A large number of variants of human serum albumin have been described.
 C:Genetics:
 A:Gene: GDB:ALB
 A:Cross-references: GDB:119990; OMIM:103600
 A:Map position: 4q11-4q13
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; py
 F.1-18/Domain: signal sequence #status predicted <SIG>
 F.19-24/Domain: propeptide #status experimental <PRO>
 F.25-609/Product: serum albumin #status experimental <MPT>
 F.29-202/Domain: serum albumin repeat homology <SA1>

F:166-174/Product: kinotensin status experimental <KIP>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA2>
F:27/Binding site: copper (Hs) #status predicted
F:77-86.99-115.114-125.148-193.192-201.224-270.269-277.289-303.302-313.340-395.394-393.4
F:214/Binding site: pyridoxal phosphate (lys) (covalent) #status experimental

Query Match 100.0%; Score 3.03; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 9.28-198; Mismatches 0; Indels 0; Gaps 0;
Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGSENFKALVLAFAQYQQCFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHKEVAHRFKDLGSENFKALVLAFAQYQQCFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLPLVRREV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLPLVRREV 144
QY 121 DVMCTAFHNEETFKKLYLYEYARRHPYFAPPELLFFPAARYKAAFTCCQAAQKAAACLLP 180
DB 145 DVMCTAFHNEETFKKLYLYEYARRHPYFAPPELLFFPAARYKAAFTCCQAAQKAAACLLP 204
QY 181 KLDELDECKKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFAEFAEYKLVTDLT 240
DB 205 KLDELDECKKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFAEFAEYKLVTDLT 264
QY 241 VHTCCHGDLLECAADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAFVENDEMPA 300
DB 265 VHTCCHGDLLECAADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAFVENDEMPA 324
QY 301 DPLSLAANDVESKQKVAEAKQVFLGMFLYEAARRHPDYSVLLLLAKTYETTLK 360
DB 325 DPLSLAANDVESKQKVAEAKQVFLGMFLYEAARRHPDYSVLLLLAKTYETTLK 384
QY 361 CAADPHCEYAKVDFEFPKLVPEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
DB 385 CAADPHCEYAKVDFEFPKLVPEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 444
QY 421 PTLVEVSRNLGVKGVKCKKHPKAPCAEDYLSVNLQVLVHEKTPVSRVTKCTES 480
DB 445 PTLVEVSRNLGVKGVKCKKHPKAPCAEDYLSVNLQVLVHEKTPVSRVTKCTES 504
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQTAALVELVGHKPKAT 540
DB 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQTAALVELVGHKPKAT 564
QY 541 KEQLKAVMDDFAAFEVKCKCKADKCTCFABEGKLVLAASQAALGL 585
DB 565 KEQLKAVMDDFAAFEVKCKCKADKCTCFABEGKLVLAASQAALGL 609

RESULT 2
A47391
serum albumin precursor - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #ext_change 20-Aug-1995
C:Accession: A47391
R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, B.; Smith, D.G.; Dwulet, J.; Putnam, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in biliary
A:Reference number: A47391; MUID:93211971; PMID:8460152
A:Contents: B/B homozygote
A:Accession: A47391
A>Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-600 <MAT>
A:Cross-references: GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G342295
A:Experimental source: liver
A>Note: sequence extracted from NCBI backbone (NCBI:128282, NCBI:128281)
C:Superfamily: serum albumin; serum albumin repeat homology
F:21-194/Domain: serum albumin repeat homology <SA1>
F:213-386/Domain: serum albumin repeat homology <SA2>

F:405-584/Domain: serum albumin repeat homology <SA3>
Query Match 94.8%; Score 2942; DB 2; Length 600;
Best Local Similarity 93.5%; Pred. No. 48-187;
Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGSENFKALVLAFAQYQQCFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 17 DTKSEVAHRFKDLGSENFKALVLAFAQYQQCFEDHVKLVNEVTEFAKTCVADESAAE 76
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLPLVRREV 120
DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLPLVRREV 136
QY 121 DVMCTAFHNEETFKKLYLYEYARRHPYFAPPELLFFPAARYKAAFTCCQAAQKAAACLLP 180
DB 137 DVMCTAFHNEETFKKLYLYEYARRHPYFAPPELLFFPAARYKAAFTCCQAAQKAAACLLP 196
QY 181 KLDELDECKKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFAEFAEYKLVTDLT 240
DB 197 KLDELDECKKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFAEFAEYKLVTDLT 256
QY 241 VHTCCHGDLLECAADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAFVENDEMPA 300
DB 257 VHTCCHGDLLECAADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAFVENDEMPA 316
QY 301 DPLSLAANDVESKQKVAEAKQVFLGMFLYEAARRHPDYSVLLLLAKTYETTLK 360
DB 317 DPLSLAANDVESKQKVAEAKQVFLGMFLYEAARRHPDYSVLLLLAKTYETTLK 376
QY 361 CAADPHCEYAKVDFEFPKLVPEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
DB 377 CAADPHCEYAKVDFEFPKLVPEPQNLKQNCLEFQELGEYKFNALLVRYTKVPQVST 436
QY 421 PTLVEVSRNLGVKGVKCKKHPKAPCAEDYLSVNLQVLVHEKTPVSRVTKCTES 480
DB 437 PTLVEVSRNLGVKGVKCKKHPKAPCAEDYLSVNLQVLVHEKTPVSRVTKCTES 496
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQTAALVELVGHKPKAT 540
DB 497 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQTAALVELVGHKPKAT 556
QY 541 KEQLKAVMDDFAAFEVKCKCKADKCTCFABEGKLVLAASQAAL 583
DB 557 KEQLKAVMDDFAAFEVKCKCKADKCTCFABEGKLVLAASQAAL 599

RESULT 3
S57632
serum albumin precursor - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #ext_change 20-Aug-1999
C:Accession: JC4660; S57632
R:Halper, C.; Grigori, F.; Hentges, F.
Gene 189, 295-296, 1996
A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
A:Reference number: JC4660; MUID:96194824; PMID:8647469
A:Accession: JC4660
A:Molecule type: mRNA
A:Residues: 1-608 <H12>
A:Cross-references: EMBL:X84842; NID:G886484; PIDN:CAA59279.1; PID:G886485
A:Experimental source: liver
A:Comment: This protein is the major protein component in plasma. It functions as a
e in has 35 conserved cysteine residues.
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: liver; plasma
F:1-18/Domain: signal sequence #status predicted <S1>
F:19-24/Domain: propeptide #status predicted <PRP>
F:25-608/Product: serum albumin #status predicted <MAT>
F:25-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>

Query Match 84.4%; Score 2623; DB 2; Length 608;
Best Local Similarity 82.0%; Pred. No. 8, 2e-166;
Matches 478; Conservative 52; Mismatches 53; Indels 0; Gaps 0;

QY 1 DAHKEVAHFRKDLGEBNFKALVLIAPQYQQCPFECHVKLVNVTFAKTCVADESAAE 60
DB 25 EAHQSEIAHFRNDLGEHFRGVLVAFSQYQQCPFECHVKLVNVTFAKTCVADESAAE 84
QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFQHKDNDPNLPLRVPEV 120
DB 85 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFQHKDNDPNLPLRVPEV 144
QY 121 DVMCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLP 180
DB 145 DAMCTAFHNEEQFLGKYLYEIAARRHPYFYAPPELLFYAEEYKGVTFECCQAADKAACLP 204
QY 181 KLDELDEGKASSAKORLKCASLOKQEGEAFKAWAVARLSQRPFAEFAEVSCLVDTLTK 240
DB 205 KVDALREKVLASSAKERLKCASLOKQEGEAFKAWAVARLSQRPFAEFAEVSCLVDTLTK 264
QY 241 VHTCCCHGDLLECAADDRADLAKYICENQDSISSKLECCCKPILKSHCIAEVENDEMPA 300
DB 265 IHKECCCHGDLLECAADDRADLAKYICENQDSISSKLECCCKPILKSHCIAEVENDEMPA 324
QY 301 DLPLSAAADVFESKDVCKNYAEAKDVFLGMFLYFYARRHPDYVSVLLRLAKTYETITLKC 360
DB 325 DPLPLAVDEVEDKEVCNKYQAEAKDVFLGTLFYYSRRHPDYVSVLLRLAKTYETITLKC 384
QY 361 CAADDPHECYAKVDFEKLVEEPQNLIKONCELFQELGEYKFNALLVRYTKVQVST 420
DB 385 CATDDPPACVAVHDFEKLVEEPQNLIKONCELFQELGEYKFNALLVRYTKVQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRCFSALEVDYETVYKPEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540
DB 505 LVNRRCFSALEVDYETVYKPEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 564
QY 541 KEQLKAWDDFAAFVCKCKADDDKETCFABEGKLVAAASQAL 583
DB 565 EEQLKTVMGDFGSDVKCCAAEDKEACFAEBSKLVAAASQAL 607

RESULT 4

ABOS
serum albumin precursor - horse
C:Species: Equus caballus (domestic horse)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S34053
R:Ho, J.X.; Holowachuk, E.W.; Norton, E.C.; Twigg, F.D.; Carter, D.C.
Eur. J. Biochem. 215, 205-212, 1993
A:Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm
A:Reference number: S34053; PMID:93345495; PMID:8344282
A:Accession: S34053
A:Molecule type: mRNA
A:Residues: 1-607 <HOA>
A:Cross-references: GB:X74045; NID:G399671; PIDN:CAAS2194.1; PID:G399672
C:Comment: Serum albumin is synthesized in the liver as prealbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membrane)
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: serum albumin #status predicted <YAT>
F:29-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status predicted
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F:263/Binding site: bilirubin (Iys) #status predicted

Query Match 79.8%; Score 2475.5; DB 1; Length 607;
Best Local Similarity 76.3%; Pred. No. 3e-156;
Matches 445; Conservative 69; Mismatches 68; Indels 1; Gaps 1;

QY 1 DAHKEVAHFRKDLGEBNFKALVLIAPQYQQCPFECHVKLVNVTFAKTCVADESAAE 60
DB 25 DTHKSEIAHFRNDLGEHFRGVLVAFSQYQQCPFECHVKLVNVTFAKTCVADESAAE 84
QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFQHKDNDPNLPLRVPEV 120
DB 85 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFQHKDNDPNLPLRVPEV 143
QY 121 DVMCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLP 180
DB 144 DAQCAAFQEDPDKFLGKYLYEVARRHPYFYAPPELLPHAEYKADFTTECCQAADKAACLP 203
QY 181 KLDELDEGKASSAKORLKCASLOKQEGEAFKAWAVARLSQRPFAEFAEVSCLVDTLTK 240
DB 204 KLDALKERILLSSAKERLKCASSFQNGEAFKAWAVARLSQRPFAEFAEVSCLVDTLTK 263
QY 241 VHTCCCHGDLLECAADDRADLAKYICENQDSISSKLECCCKPILKSHCIAEVENDEMPA 300
DB 264 VHKCCCHGDLLECAADDRADLAKYICENQDSISSKLECCCKPILKSHCIAEVENDEMPA 323
QY 301 DLPLSAAADVFESKDVCKNYAEAKDVFLGMFLYFYARRHPDYVSVLLRLAKTYETITLKC 360
DB 324 DPLAALADFAEDKEICXHYKDAKDVFLGTLFYYSRRHPDYVSVLLRLAKTYETITLKC 383
QY 361 CAADDPHECYAKVDFEKLVEEPQNLIKONCELFQELGEYKFNALLVRYTKVQVST 420
DB 384 CAADDPACVAVHDFEKLVEEPQNLIKONCELFQELGEYKFNALLVRYTKVQVST 443
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 444 PTLVEIGRTJGVGSRCKLPESERLPCSENHIALALNRLCVLHEKTPVSEKIKCCTDS 503
QY 481 LVNRRCFSALEVDYETVYKPEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540
DB 504 LAERRPCFSALEVDYETVYKPEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 563
QY 541 KEQLKAWDDFAAFVCKCKADDDKETCFABEGKLVAAASQAL 583
DB 564 KEQLKTVJGNFSAPFAKCCGREDKEACFAEBSKLVAAASQAL 606

RESULT 5
ABOS
serum albumin precursor [validated]; - bovine
N:Alternate names: 67K protein; prealbumin
C:Species: Bos primigenius taurus (cattle)
C>Date: 24-Apr-1984 #sequence_revision 30-Sep-1993 #text_change 18-Aug-2000
C:Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A92309; A31458;
R:Holowachuk, E.W.; Stoltzenberg, J.K.; Reed, R.G.; Peters Jr., T.
Submitted to the EMBL Data Library, August 1991
A:Description: Bovine serum albumin: cDNA sequence and expression.
A:Reference number: A38885
A:Accession: A38885
A:Molecule type: mRNA
A:Residues: 1-607 <HOA>
A:Cross-references: EMBL:M73215
R:Hiyama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.
Biochem. Biophys. Res. Commun. 173, 639-646, 1990
Eur. J. Biochem. 98, 477-485, 1979
A:Title: Rapid confirmation and revision of the primary structure of bovine serum a)
A:Reference number: A36401; PMID:91083649; PMID:2260975
A:Accession: A36401
A:Molecule type: protein
A:Residues: 25-41, 'H', 43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HIR>
R:MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.
Eur. J. Biochem. 98, 477-485, 1979
A:Title: Biosynthesis of bovine plasma proteins in a cell-free system.
A:Reference number: A91258; PMID:80024278; PMID:488109
A:Accession: A91258
A:Molecule type: protein

A:Residues: 1-32 <YAG>
 R.Hsieh, J.C.; Lin, F.P.; Tam, M.F.
 Anal. Biochem. 170, 1-9, 1988
 A:Title: Electrophoretic transfer of an analytical isoelectrofocusing gel
 A:Reference number: A60808; MUID:88267456; PMID:3389500
 A:Accession: B60808
 A:Molecule type: protein
 A:Residues: 25-41 <HS1>
 R.Strawich, E.; Glanville, M.J.
 Eur. J. Biochem. 191, 47-56, 1990
 A:Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is also
 A:Reference number: S13780; MUID:90336641; PMID:2379503
 A:Accession: S10780
 A:Molecule type: protein
 A:Residues: 25-41, 43-57, 59-64 <STR>
 R.Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
 J. Immunol. 143, 1680-1694, 1989
 A:Title: Structures of histamine-releasing peptides formed by the action of acid proteases
 A:Reference number: A45800; MUID:89341406; PMID:2474609
 A:Accession: D45800
 A:Molecule type: protein
 A:Residues: 163-172 <CAR>
 R.Carraway, R.E.; Mitra, S.P.; Cochrane, D.E.
 J. Biol. Chem. 262, 5968-5973, 1987
 A:Title: Structure of a biologically active neurotensin-related peptide obtained from pe
 A:Reference number: A26693; MUID:87134805; PMID:243711
 A:Accession: A26693
 A:Molecule type: protein
 A:Residues: 165-172, 175 <CA2>
 R.Reed, R.G.; Putnam, F.W.; Peters Jr., T.
 Biochem. J. 191, 867-868, 1980
 A:Title: Sequence of residues 400-403 of bovine serum albumin.
 A:Reference number: A90309; MUID:82023364; PMID:7283978
 A:Accession: A90309
 A:Molecule type: protein
 A:Residues: 402-433 <R3E>
 R.Brown, J.R.
 Fed. Proc. 34, 591, 1975
 A:Title: Structure of bovine serum albumin.
 A:Reference number: A94551
 A:Accession: A94551
 A:Molecule type: protein
 A:Residues: 190-195 <BR2>
 R.Brown, J.R.
 Fed. Proc. 33, 1389, 1974
 A:Reference number: A91457
 A:Contents: annotation; disulfide bonds
 R.Werlen, R.C.; Offord, R.E.; Rose, K.
 Biochem. J. 302, 907-911, 1994
 A:Title: Preparation and characterization of novel substrates of insulin proteinase (EC
 A:Reference number: S55232; MUID:95031935; PMID:7945219
 A:Accession: S55232
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 529-536, 569-572 <WER>
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; copper binding; duplication; plasma
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-24/Domain: propeptide #status experimental <PRO>
 F:25-607/Product: serum albumin #status experimental <MPT>
 F:29-201/Domain: serum albumin repeat homology <SA1>
 F:220-393/Domain: serum albumin repeat homology <SA2>
 F:412-591/Domain: serum albumin repeat homology <SA3>
 F:27/Binding site: copper (His) #status predicted
 F:77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392, 4

Query Match

78.8%; Score 2446.5; DB 1; Length 607;

Best Local Similarity 75.6%; Pred. No. 2.5e-154;
 Matches 441; Conservative 71; Mismatches 70; Indels 1; Gaps 1;
 QY 1 DAHKSEVAHRFKDLGEENFKALVLAFAQYLOOCFEDHVKLVREVTFAKTCVADESAE 60
 DB 25 DTHKSEIAHRFKDLGEBOPKGLVLAFAQYLOOCFDEHVKLVNELTEFAKTCVADESHA 84
 QY 61 NDKSLHTLFGSKLCTVATLRETYGEMADCAKQPERNECFLOHKDDNPNLPRVLRREV 120
 DB 85 GGEKSJHTLFGDLCKVASLRETYGEMADCCQPERNECFSLSHKDDSPDLPKL-KEDP 143
 QY 121 DVNCTAFHNEETFLKKYVEIARRHPYFAZELLFFAKRYKAAFTCCCAADKAAACLLP 180
 DB 144 NLCDEFKADKKFWCKYLYEARRHPYFAZELLFYKYNKNGVFQDCQCAEDKACALLP 203
 QY 181 KUJELRDEGKASAKORLKASLQKGERAFKAMAVARLSQFPKAEFAEYVKLVTDLTK 240
 DB 204 KETMREKVLASSAROLKASIQKFGERALKAMSVARLSQFPKAEFVEVTKVTDLTK 263
 QY 241 VHTCCGGLLECAODRADLAKYICENODSISKKKECEKPLLEKSHCIAVENDEMPA 300
 DB 264 VHKCCGGLLECAODRADLAKYICONQDTISKKKECCPKLEKSHCIAEVEKDA-PE 323
 QY 301 DLPSLAADFEVSKDVCNKYAEAKDFLGMFLYVEYARRHPDYVWLLRLAKTYETTLKRC 360
 DB 324 NUPPLTADFAEDKDKVCKNYQEAQDAFLGSLFYEYSRRHPZYAVSVLLRLAKTYETTLKRC 383
 QY 361 CAAADPHECIYAKVDFEKLVEEPQNLKQNCLEFQELGEYKFNALVRYTKVPQVST 420
 DB 384 CAKDDPHACYSTVDFKLHVLDEPQNLKQNCLEFQELGEYKFNALVRYTKVPQVST 443
 QY 421 PTLVEVSRNLGKVGSKCKKHPKAPCAEDVSVVNLGLCVLHEKTPVSDRVTKCCTES 480
 DB 444 PTLVEVSRNLGKVGSKCKKHPKAPCAEDVSVVNLGLCVLHEKTPVSDRVTKCCTES 503
 QY 481 LVNRPCFSALEVDETYVPKFNATFTFHADICTLSEKERQIKQTALVELVKKHKPAT 540
 DB 504 LVNRPCFSALEVDETYVPKFNATFTFHADICTLSEKERQIKQTALVELVKKHKPAT 563
 QY 541 KEOLKAWDCFAAFVEKCKCKADKTCFAEEGKULVAASQAAL 583
 DB 564 EEOELKAWDCFAAFVEKCKCKADKTCFAEEGKULVAASQAAL 606
 RESULT 6
 AESS
 serum albumin precursor - sheep
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C:Accession: S06936
 R:Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.
 Nucleic Acids Res. 17, 10495, 1989
 A:Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
 A:Reference number: S06936; MUID:90098888; PMID:2602160
 A:Accession: S06936
 A:Molecule type: mRNA
 A:Residues: 1-607 <BRO>
 A:Cross-references: EMBL:X17055; NID:G1386; PID:CAA34903.1; PID:G1387
 C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper
 and iron, and promotes the transfer of these elements across the men
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; metal binding; plasma
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-24/Domain: propeptide #status predicted <PRO>
 F:25-607/Product: serum albumin #status predicted <MAT>
 F:29-201/Domain: serum albumin repeat homology <SA1>
 F:220-393/Domain: serum albumin repeat homology <SA2>
 F:412-591/Domain: serum albumin repeat homology <SA3>
 F:27/Binding site: copper (His) #status predicted
 F:77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-393
 F:263/Binding site: bilirubin (Lys) #status predicted

Query Match

78.4%; Score 2432.5; DB 1; Length 607;


```
Best Local Similarity 75.0%; Pred. No. 2.1e-153;
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

QY 1 DAHSEVAHRRFKOLGEENPKALVLIAPACYQQCPFEHVKLVNEVTEFAKTCVADSEAE 60
DB 25 DTHKSEIAHRRNDLGEENFQGLVLIAPSYQYQQCPFEHVKLVNKLTEFAKTCVADSEAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVLRPEV 120
DB 85 GCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVLRPEV 143
QY 121 DVMTAFPHDNEETPLKKYLYEIAIRRHYPFYAPPELLFFAKYKAAFTSCCAADKAAACLLP 180
DB 144 DTLCAEPKADKKFKGKYLVEYARRHPFYAPPELLFYANKYNGVFOECCQAECKGACLLP 203
QY 131 KJDELDRDEGKASSAKORLKCASLQKQGERAFKAWAVARLSQRFKAEFAEYSKLVITLTK 240
DB 204 KIDAMREKVLASSARQURCASIQGERALKAMSVARLSQRFKAEFAEYSKLVITLTK 263
QY 241 VHTECCCHGDLLECCADDRADLAKYICENODSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 264 VHKECCHGDLLECCADDRADLAKYICDHQALSSKLEKCEKPLLEKSHCIAEVENDEMPA 323
QY 301 DLPSLAADPFVESKDVCKYAKAKQVFLGMFLYEVARRHPDYSVVLRLAKTYETTLEK 360
DB 324 NPLPLTADFAEDKEVKCKYAKAKQVFLGMFLYEVARRHPDYSVVLRLAKTYETTLEK 383
QY 361 CAADPHCEKAVKDFEKPVLVEPQNLIKONCELFEQOLGEYKFNALLVRYTKKQPOVST 420
DB 384 CAKEDPHACATVDFPKLKHLDVDEPNLIKONCELFEKGEYGFQNALIVRYTKKQPOVST 443
QY 421 PLVLESYRNGLKVGSKCKCPKAPCAEDYLSVLNQCVLREKTPVSRVTKKCTES 480
DB 444 PLVLESYRNGLKVGSKCKCPKAPCAEDYLSVLNQCVLREKTPVSRVTKKCTES 503
QY 481 LVNRRPCFSALVDETVYPKFNAEFTTHADICTLSEKERQIKQOTALVELVKKKPKAT 540
DB 504 LVNRRPCFSALVDETVYPKFNAEFTTHADICTLSEKERQIKQOTALVELVKKKPKAT 563
QY 541 KEQLKAVMDFFAFAVEKCKCKADQKTCFAEKGKLVAAASQAL 583
DB 564 DEQLKAVMDFFAFAVEKCKCKADQKTCFAEKGKLVAAASQAL 606

RESULT 7
ABR'S
serum albumin precursor - rat
N:Alternate names: prealbumin
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-May-1979 #sequence revision 31-May-1979 #text change 22-Jun-1999
C:Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233
R:Sargent, T.D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981.
A:Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A:Reference number: A93872; MUID:81223722; PMID:7317712
A:Accession: A93872
A:Molecule type: mRNA
A:Residues: 1-608 <STR>
A:Cross-references: GB:J00698; NID:955627; PIDN:CAA24532.1; PID:955628
R:Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rockey, J.A.; Alberts, A.W.
J. Biol. Chem. 252, 6846-6855, 1977
A:Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analysis
A:Reference number: A92211; MUID:77249657; PMID:893447
A:Note: Cleavages during protein maturation
A:Accession: A92211
A:Molecule type: protein
A:Residues: 1-38 <STR>
R:Isemura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A:Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage
A:Reference number: A91946; MUID:78109429; PMID:564345
A:Accession: A91946
A:Molecule type: protein
```

QY 421 PTLVEVSRLNGKSGKCKKHPKAPKAPCAEDYLSWVNLQVLVHEKTPVSDRVTKCTES 490
 DB 445 PTLVEAARNLGRVGTCKCTLPFAQRLPOVEDYLSAILNRLCVLHEKTPVSEKVTKCCSGS 504
 QY 491 LVNRRPCFSALVEDETYVPKPEFNAETTFHADICTLSEKERQKKQTALVELVKKHPKAT 540
 DB 505 LVRRPCFSALTVDYVPKPEFNAETTFHSDICTLSEKERQKKQTALVELVKKHPKAT 564
 QY 541 KEQLKAVVDFAAFVVEKCKKADDETCFABEGKKLVAASQAL 583
 DB 565 EDQLKTVMGDFQAFVDRCKKADNDNFATFEGPNLVASKEAL 607
 RESULT 8
 ABPGS
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C:Accession: S01382; A61066
 R:Weinstock, J.; Baldwin, G.S.
 Nucleic Acids Res. 16, 9045, 1988
 A:Title: Nucleotide sequence of porcine liver albumin.
 A:Reference number: S01382; XUID:89C:6582; PMID:3174440
 A:Accession: S01382
 A>Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-605 <WEI>
 A:Cross-references: EMBL:X12422; NID:G1875; PID:CAA30970.1; PID:G833798
 R:Lineback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.
 J. Bone Miner. Res. 4, 235-241, 1989
 A:Title: Serum albumin and its acid hydrolysis: peptides dominate preparations of mineral
 A:Reference number: A61006; MUID:89269769; PMID:2728927
 A:Accession: A61006
 A:Molecule type: protein
 A:Residues: 23-51,'X',53-54,'XXXGY',146,'E',148,'E',150-151,'XVN',155 <LM>
 A:Experimental source: dental enamel
 A:Note: albumin and other serum proteins are also found in bone
 A:Comment: Serum albumin is synthesized in the liver as prealbumin. It binds copper,
 steroid hormones (weak bonds with these hormones promote their transfer across the membra
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; metal binding; plasma
 F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>
 F:17-22/Domain: propeptide #status predicted <PSO>
 F:23-605/product: serum albumin #status predicted <MAT>
 F:27-199/Domain: serum albumin repeat homology <SA1>
 F:218-391/Domain: serum albumin repeat homology <SA2>
 F:410-589/Domain: serum albumin repeat homology <SA3>
 F:75-84,97-113,112-123,145-190,189-198,221-267,266-274,296-300,299-310,337-382,381-390,4
 F:261/Binding site: bilirubin (lys) #status predicted
 Query Match 77.7%; Score 2411.5; DB 1; Length 605;
 Best Local Similarity 76.0%; Pred. No. 5e-152;
 Matches 438; Conservative 67; Mismatches 70; Indels 1; Gaps 1;
 QY 1 DAHSEVAHRFKDLGEEFNKALVLIAPAYLQOCFFEDHYKLVNEVTEFAKTCVADESA 60
 DB 23 DTYKSEIAHRFKDLGEEFNKALVLIAPAYLQOCFFEDHYKLVNEVTEFAKTCVADESA 82
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNNPLRLVRPEV 120
 DB 83 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNNPLRLVRPEV 141
 QY 121 DVMTCTAFHNEETFKKLYLYEYIARRHPYFAPPELLFAKRYKAAFTCCCAAKAAILP 180
 DB 142 VALCADFOEDQKFGKGYEYIARRHPYFAPPELLFAKRYKAAFTCCCAAKAAILP 211
 QY 181 KLDELDEGKASSAKORLKASLOKFGGERAFKAWAVARLSORFPKAEVSKLVDTLTK 240
 DB 202 KIEHLREKVLTSAAKORLKASLOKFGGERAFKAWAVARLSORFPKAEVSKLVDTLTK 261
 QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKCECKPDLKSHCIAEVENDEMPA 300

DB 262 VHKCCHGDLLECCADRADLAKYICENQDSISSKLKCECKPDLKSHCIAEAKRDLP 321
 QY 301 DLSLAADPVESKDVCKQYAEAKDVFLGMLFYBVARRRHPDYSVVLLRLAKTYETTTLEK 360
 DB 322 DLNPFJEHDFVEDNEVKQYAEAKDVFLGMLFYBVARRRHPDYSVVLLRLAKTYETTTLEK 381
 QY 361 CAADPHSCYAKVDFDEPKLIVBEPQNLKONCELFBSOLGEYKFNALVRYTKVQVST 420
 DB 382 CAKEDDPACVATVDFKQPLVDEPKNLKONCELFBSOLGEYKFNALVRYTKVQVST 441
 QY 421 PTLVEVSRLNGKSGKCKKHPKAPKAPCAEDYLSWVNLQVLVHEKTPVSDRVTKCTES 480
 DB 442 PTLVEVARKLGVGRCKKAPBBERISCADYLSLVNRLCVLHEKTPVSEKVTKCTES 501
 QY 491 LVNRRPCFSALVEDETYVPKPEFNAETTFHADICTLSEKERQKKQTALVELVKKHPKAT 540
 DB 502 LVNRRPCFSALTVDYVPKPEFNAETTFHADICTLSEKERQKKQTALVELVKKHPKAT 561
 QY 541 KEQLKAVVDFAAFVVEKCKKADDETCFABEGKKLVAASQAL 576
 DB 562 EQLRTVLGNFAAFVQKCAAPDHEACFAVEGPKFV 597
 RESULT 9
 C5836
 albumin - Mongolian jird
 C:Species: Meriones unguiculatus (Mongolian jird)
 C:Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 19-May-2000
 C:Accession: J05838
 R:Yoshida, K.; Seto-Ohshima, A.; Sinohara, H.
 DNA Res. 4, 351-354, 1997
 A:Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in
 A:Reference number: J05838; MUID:98116663; PMID:9455485
 A:Accession: J05838
 A:Molecule type: mRNA
 A:Residues: 1-609 <YOS>
 A:Cross-references: DDBJ:AB006197; NID:G2317277; PID:BAA21765.1; PID:G2317278
 A:Experimental source: liver
 A:Superfamily: serum albumin; serum albumin repeat homology
 F:222-395/Domain: serum albumin repeat homology <SA2>
 Query Match 76.9%; Score 2387; DB 2; Length 609;
 Best Local Similarity 73.9%; Pred. No. 2.1e-150;
 Matches 430; Conservative 65; Mismatches 87; Indels 0; Gaps 0;
 QY 2 AHKSEVAHRFKDLGEEFNKALVLIAPAYLQOCFFEDHYKLVNEVTEFAKTCVADESA 61
 DB 27 AHKSEIAHRFKDLGEEFNKALVLIAPAYLQOCFFEDHYKLVNEVTEFAKTCVADESA 86
 QY 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNNPLRLVRPEV 121
 DB 87 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNNPLRLVRPEV 146
 QY 122 VMTCTAFHNEETFKKLYLYEYIARRHPYFAPPELLFAKRYKAAFTCCCAAKAAILP 181
 DB 147 AMCTAFQENASAFHGYLHEVARRHPYFYGPELLYACKYTAVALTECCADDDGACCTPK 206
 QY 182 LDELDEGKASSAKORLKASLOKFGGERAFKAWAVARLSORFPKAEVSKLVDTLTKV 241
 DB 207 LDALKEKALVSARQRLKCSMKKFGGERAFKAWAVARLSOTFFNADPAETIKLATLTKV 266
 QY 242 HTECHGDLLECCADRADLAKYICENQDSISSKLKCECKPDLKSHCIAEVENDEMPAD 301
 DB 267 TOECCHGDLLECCADRADLAKYICENQDSISSKLKCECKPDLKSHCIAEVENDEMPAD 326
 QY 302 LPSLAADPVESKDVCKQYAEAKDVFLGMLFYBVARRRHPDYSVVLLRLAKTYETTTLEK 361
 DB 327 LPALTADEVEDKQYAEAKDVFLGMLFYBVARRRHPDYSVVLLRLAKTYETTTLEK 386
 QY 362 AAADPHCYAKVDFDEPKLIVBEPQNLKONCELFBSOLGEYKFNALVRYTKVQVST 421
 DB 387 AADPHCYAKVDFDEPKLIVBEPQNLKONCELFBSOLGEYKFNALVRYTKVQVST 446


```

Db      330 LPSLVKZYEDKEVCKSPFAGHDAPMAEFVVEYSSRRHPEFSIQLWRIAKGVESLLEKCC 389
QY      362 AAADPHCYAKVDFDEKPLVVEBPQN::KNCSEJFQGLGEYKFNQALLVRYTKVPQVGP 421
Db      390 KTDNPAECYANAQEQUNQHIKETQDVVKNQCNLLHGHGEADFKSITLRYTKVPQVPTD 449
QY      422 TLVEVSRLNGKVGSKCKPEAKRMPCAEDY::SVVLNQLCVLHEKTPVSDRVTKCCTESL 481
Db      450 LLLETGKMTTITKCCQCGEDRRMACSEGYLSIVHDTCRQETTPINDNVQCCSLY 509
QY      482 VNRPPCFSALEVDYVYKFEKNAETFTFHAD::CTLSEKEROIKKQTAVALVKKHKKPKTK 541
Db      510 ANRRPCFTAMGVDTKYVPPFPDMESPDEKLCSPAEEBEVGMKLLNLJLKKRQPMTE 569
QY      542 EQKAVMDPFAFVVEKCKKADKQKTCFAESGKKLVAAASQAALGL 585
Db      570 EQIKTIADQFTAMVDRCKCKQSQDINTCFEGEGANLVQSRATLGI 633

RESULT 12
JC4258
alpha-fetoprotein precursor - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C:Date: 27-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 20-Aug-1999
C:Accession: JC4258
R:Nishio, H.; Gibbs, P.E.M.; Vinghetti, P.P.; Zielinski, R.; Dugaiczky, A.
Gene 182, 213-220, 1995
A:Title: The chimpanzee alpha-fetoprotein-encoding gene shows structural similarity to b
A:Reference number: JC4258; MUID:96032345; PMID:7557431
A:Accession: JC4258
A:Molecule type: DNA
A:Residues: 1-609 <NIS>
A:Cross-references: GS:U2196; NID:984131; PID:AAA0164.1; P.D:8841312
C:Comment: This protein is a plasma protein produced in the fetal and neonatal liver and
c similar properties and structure.
C:Genetics:
A:Gene: afp
A:Map position: 3p
A:Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: glycoprotein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-609/Product: alpha-fetoprotein #status predicted <MAT>
F:28-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:42,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.4%; Score 1253.5; DB 2; Length 609;
Best Local Similarity 40.4%; Pred. No. 2e-75;
Matches 236; Conservative 115; Mismatches 231; Indels 7; Gaps 3;

QY      3 HKSE-----VAHFKLIGENFVALVLIAPQV::QQCPEDHVKLVNTEFAKTCVADE 57
Db      22 HRNEYGIASLDYQCTAEINLTDJAT::FFAQVQVQATYKEVSKYKQALTA::EKPTGDE 81
QY      58 SAENCKSKJLT::FGCKCTVATIRETYGEMADCCAKQEPERNECFLOHKDCKP::NLPLRV 116
Db      82 QSAGCLNQJPAFLEE::CREKEILKVGH::SDCCSQSEGRHNCFLAHKKPTFASIPFC 140
QY      117 RPEVDVNTAFHNEETFLKKYLYEARRHPYVYAPELLFAKRYKAAATECCQAADKAA 176
Db      141 VPEPVTSCAEYEDRETFMKNF::YEIARRHPF::YAPT::ILWAARYSKTIPSCKAENAVE 200
QY      177 CLLPKJLDELDEGKASSAKORLKAS::CKGGERAFKAWAVAR::SQRPFAEAEVSKLVT 236
Db      201 CFQTKAATVTKELRESSLNQHACAVNKNFSTRTFQA::TVTKLSQKFT::KVNFTTEIKLYJ 260
QY      237 DLTQVHTECCHGDLLECADRAD::LAKYICENQDSISKLECCCEKPLLEKSHKICAEVND 296
Db      261 DVAVHVEHCCGVDLQDGEKIMSYICQQDTLSNKTTECCKLT::LGRGQCLIHAKND 320
QY      297 EMPADPLSLAADC FVESKDVCKNVAEAKDVFGLGMLFVEYARRHPYVSVVJLLRLAKTYET 356

```

```

Db      321 EKPEG::SPN::NRFGLGRDNQFSSGEKNIFLASFVHEYSRRHPQLAVSVILRAVKYQEL 380
QY      357 LEKCCAAAPHECYAKVDFDEKPLVVEBPQNLIKONCELFQGLGEYKFNQALLVRYTKVP 416
Db      381 LEKCFOTENPLEQDQGEELQKYIQESQALAKRSCGLFQKLGVEYVLQNAF::VAVTKKAP 440
QY      417 QVSTPTLVEVSRLNGKVGSKCKPEAKRMPCAEDY::SVVLNQLCVLHEKTPVSDRVTKC 476
Db      441 QJSSSEJMAITRQMAATAAT::CCOLSEDKLLACGEGAAD::IGHLCIRHETTPVNPFGVQC 500
QY      477 CTESLVNRPPCFSALEVDYVYKFEKNAETFTFHAD::CTLSEKEROIKKQTAVALVKKH 536
Db      501 CTSYANRRPCFTSSLVVDYVYPPAFSDDKF::FHKDLCOAGVALQTKQBEF::JNLVKQK 560
QY      537 PKATKQLKAVMDPFAFVVEKCKKADKQKTCFAESGKKLVAAASQAALGL 585
Db      561 PQTESQLEAV::ADFSGLLEKCCQCGEQBEVCFEAESGQK::SKTRAAALGV 609

RESULT 13
FPHU
alpha-fetoprotein precursor [validated] - human
N:Alternate names: AFP; alpha-1-fetoprotein; alpha-fetoglobulin
C:Species: Homo sapiens (man)
C:Date: 19-Feb-1984 #sequence_revision 19-Feb-1984 #text_change 08-Dec-2000
C:Accession: A26824; S37655; A93961; A91497; A23699; A61480; A90624; A90757; A93042;
R:Gibbs, P.E.M.; Zielinski, R.; Boyd, C.; Dugaiczky, A.
Biochemistry 26, 1332-1343, 1987
A:Title: Structure, polymorphism, and novel repeated DNA elements revealed by a comp
A:Reference number: A26824; MUID:87185438; PMID:2436661
A:Accession: A26824
A:Molecule type: DNA
A:Residues: 1-609 <GIS>
A:Cross-references: GS:X16110; NID:G773678; PIDN:AAB38754.1; PID:gl79236
R:McVey, J.H.; Michaelides, K.; Hansen, L.P.; Ferguson-Smith, M.; Tilghman, S.; Krum
Hum. Mol. Genet. 2, 379-384, 1993
A:Title: A G-A substitution in an HNF I binding site in the human alpha-fetoprotein
A:Reference number: S37655; MUID:93278385; PMID:7684942
A:Accession: S37655
A:Molecule type: DNA
A:Residues: 1-28 <MCV>
A:Cross-references: EVBL:Z19532; NID:G28527; PIDN:CAA79592.1; PID:G28528
A:Note: The authors translated the codon TAT for residue 26 as Thr
R:Morinaga, T.; Sakai, M.; Wegmann, T.G.; Tamaoki, T.
Proc. Natl. Acad. Sci. U.S.A. 80, 4604-4608, 1983
A:Title: Primary structures of human alpha-fetoprotein and its mRNA.
A:Reference number: A93961; MUID:83273664; PMID:6192439
A:Accession: A93961
A:Molecule type: mRNA
A:Residues: 1-609 <MOR>
A:Cross-references: GS:J00077; NID:G311348; PIDN:CAA24758.1; PID:G31351
R:Beattie, W.S.; Dugaiczky, A.
Gene 20, 415-422, 1982
A:Title: Structure and evolution of human alpha-fetoprotein deduced from partial seq
A:Reference number: A91497; MUID:83158778; PMID:6187626
A:Accession: A91497
A:Molecule type: mRNA
A:Residues: 428-556 <BEA>
A:Cross-references: GS:J00076
R:Fucci, P.; Siciliano, R.; Malorni, A.; Marino, G.; Tecce, M.F.; Ceccarini, C.; Ter
Biochemistry 30, 5061-5066, 1991
A:Title: Human alpha-fetoprotein primary structure: a mass spectrometric study.
A:Reference number: A23699; MUID:91242409; PMID:1709810
A:Accession: A23699
A:Molecule type: protein
A:Residues: 19-45:60-97:102-107:122-184:187-249:255-489:507-609 <FUC>
R:Tecce, M.F.; Terrana, B.; Giuliani, M.M.; Ceccarini, C.
J. Nucl. Med. Allied Sci. 34, 213-216, 1990
A:Title: Characterization of in vitro expressed human alpha-fetoprotein as highly rei
A:Reference number: A61480; MUID:91255826; PMID:1709209
A:Accession: A61480
A:Molecule type: protein

```

A:Residues: 19-45;63-97;102-137;122-184;187-249;255-489;507-609 <TEC>
 R:Yachnin, S.; Hsu, R.; Heinriksson, R.L.; Miller, J.B.
 Biochim. Biophys. Acta 493, 418-428, 1977
 A:Title: Studies on human alpha-fetoprotein. Isolation and characterization of monomeric
 A:Reference number: A90624; MUID:77242506; PMID:70228
 A:Accession: A90624
 A:Molecule type: protein
 A:Residues: 'S', 20-22, 'S', 24-35 <YAC>
 A:Note: dimeric and trimeric forms have been found in addition to the monomeric form
 R:Aoyagi, Y.; Ikenaka, T.; Ichida, F.
 Cancer Res. 37, 3663-3667, 1977
 A:Title: Comparative chemical structure of human alpha-fetoproteins from fetal serum and
 A:Reference number: A90757; MUID:78001760; PMID:71198
 A:Accession: A90757
 A:Molecule type: protein
 A:Residues: 'S', 20-30, 'A', 32-37, 'A' <AOY>
 R:Ruoslahti, E.; Pihko, H.; Vaheri, A.; Sepala, M.; Viroinen, M.; Kontinen, A.
 Johns Hopkins Med. J. Suppl. 3, 249-255, 1974
 A:Title: 20. Alpha fetoprotein: structure and expression in man and inbred mouse strains
 A:Reference number: A93042; MUID:75018719; PMID:4138095
 A:Accession: A93042
 A:Molecule type: protein
 A:Residues: 'S', 20-24, 'Q', 26-30, 'A', 32-35, 'E', 37-39 <RUC>
 R:Sakai, M.; Morinaga, T.; Urano, Y.; Watanabe, K.; Wegmann, T.G.; Tamaoki, T.
 J. Biol. Chem. 260, 5055-5060, 1985
 A:Title: The human alpha-fetoprotein gene. Sequence organization and the 5' flanking region
 A:Reference number: A92520; MUID:85192629; PMID:2580830
 A:Contents: annotation; gene, exons and introns
 R:Aoyagi, Y.; Ikenaka, T.; Ichida, F.
 Cancer Res. 39, 3483-3486, 1979
 A:Title: Copper(II)-binding ability of human alpha-fetoprotein.
 A:Reference number: A90758; MUID:79001617; PMID:83265
 A:Contents: annotation; metal binding
 R:Aoyagi, Y.; Ikenaka, T.; Ichida, F.
 Cancer Res. 39, 3574-3574, 1979
 A:Title: Alpha-fetoprotein as a carrier protein in plasma and its bilirubin-binding ability
 A:Reference number: A90759; MUID:83001710; PMID:89950
 A:Contents: annotation; bilirubin binding
 C:Comment: AFP is synthesized by the fetal liver and yolk sac. It occurs in the plasma of
 C trace amounts after birth. The serum level in adults is usually less than 40 ng/ml. AFP
 C:Comment: Human AFP binds copper, nickel, and fatty acids as well as, and the bilirubin
 C:Properties:
 C:Genetics:
 A:Gene: GDB:AFP
 A:Cross-references: GDB:119660; OMIM:104150
 A:Map position: 4q11-q13
 A:Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 435/2; 476/3; 551/3
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma
 F:1-18/Domain: signal sequence #status predicted <S>
 F:19-609/Product: alpha-fetoprotein #status experimental <MAT>
 F:29-202/Domain: serum albumin repeat homology <SAI>
 F:221-394/Domain: serum albumin repeat homology <SA2>
 F:413-592/Domain: serum albumin repeat homology <SA3>
 F:22/Binding site: copper (His) #status experimental
 F:99-114,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-472
 F:249/Binding site: bilirubin (Lys) #status predicted
 F:251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match: 40.3%; Score 1249.5; DB 1; Length 609;
 Best Local Similarity 39.9%; Pred. No. 3.8e-75;
 Matches 235; Conservative 116; Mismatches 23; Indels 7; Gaps 3;

Db 14: VPEVTSCEAYEDRETMMKFIYEARRHPPLVAPITILJWAARYDKIISSCCAKENAVE 200
 QY 177 CLLPKLDELDRDEGKASSAKORLKCAQLKQGERAFKAWAVARLSQRPFAEFAEVSILVT 236
 Db 201 CFQTKAATVTKELRESSLLNOHACAVMKNFTGTRIFQAITVTKLSCKTKYNFTIOLKLV 260
 QY 237 DLTQVHTCCGGDLJECADDRADLAKYICENODSSSKLKECCCKPPLLEKSHCHIAEVND 296
 Db 261 DVAHVHECHCRGVDLDCLQDGEKIMSYICQDDTSNKTITSCCKLTIERGQCIIHAEND 320
 QY 297 EMPADLPALADDFVESKDVCKYAEAKDVLGMFLYEVARHPDYSVVLLRLAKTYETT 356
 Db 321 EKPEGLSPNLNRFUGDDRFNQSSEKKNIFJASVHEYSRRHPQJAVSVIRVAKGQEL 380
 QY 357 LEKCAAADPHECYAKVDFEFLKPLVEEPQNLIKQNCLEFQOLGEYKFKONALLVRYTKVP 416
 Db 381 LEKCFQTEPLECCQDKGEEELQKVIQESQALAKSKGFLQKLGYYLQNAFLVATKKAP 440
 QY 417 QVSTPTLVEVSRNLGKVGSKCKHPEAKRMPCADYLSVNLQLCVLHEKTPVSDRVTKC 476
 Db 441 QLTSELMAITRKMAATAATCCQLSEDKLLACGGAADIIIGHLCIRHEMTVPNPGVQC 500
 QY 477 CTESLVNRRPCFSALVDVETVYKPEFNAETFEHADICTLSEKEROIKKOTALVELVGHK 536
 Db 501 CTSSVANNRRPCFSSLVVDSTVTPAFSDDKFIHKDLCOAGVALQTMKQELINLVKQK 560
 QY 537 PKATKEQLKAVMCDFAAFVEKCKCAADKCTCFABEGKKLVAAASQAALGL 585
 Db 561 PQITEEQLEAVIADFSGLLEKCCQGEQVCEFAEEGKLISKTRAAIGV 609
 RESULT 14
 PFCO
 A:Species: Gorilla gorilla (gorilla)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C:Accession: A37970
 R:Ryan, S.C.; Zielinski, R.; Dugaiczky, A.
 Genomics 9, 60-72, 1991
 A:Title: Structure of the gorilla alpha-fetoprotein gene and the divergence of prima
 A:Reference number: A37970; MUID:91169517; PMID:1706310
 A:Accession: A37970
 A:Molecule type: DNA
 A:Residues: 1-609 <RYA>
 A:Cross-references: GB:M38272; NID:9817963; PIDN:AAA73520.1; PID:9177041
 C:Genetics:
 A:Map position: 4q11-12
 A:Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3;
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma
 F:1-18/Domain: signal sequence #status predicted <S>
 F:19-609/Product: alpha-fetoprotein #status predicted <MAT>
 F:29-202/Domain: serum albumin repeat homology <SAI>
 F:221-394/Domain: serum albumin repeat homology <SA2>
 F:413-592/Domain: serum albumin repeat homology <SA3>
 F:22/Binding site: copper (His) #status predicted
 F:99-114,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-472
 F:249/Binding site: bilirubin (Lys) #status predicted
 F:251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match: 40.0%; Score 1242.5; DB 1; Length 609;
 Best Local Similarity 39.6%; Pred. No. 1.1e-74;
 Matches 233; Conservative 117; Mismatches 23; Indels 7; Gaps 3;

QY 3 HKSE-----VAHRFKDLGLENFKALVLAFAPYLOCCPFEDHVKLVNVEFPAKTCVADE 57
 Db 22 HRNEVGASLDYSQCTAETSLADLAIFFAQFQVETATYKEVSKWVDALTAIEKPTGDE 81
 QY 58 SAENCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNP-NLPLRV 116
 Db 82 QSSGCLNQLPAFLEELCHEKEILEKYG-SDCCSQSEGRHNCFLAHKKPTPASIFLQ 140
 QY 117 RPEVDMCTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQADKAA 176

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 15:08:38 ; Search time 25 Seconds
(without alignments)
1100.425 Million cell updates/sec

Title: US-09-833-117-18

Perfect score: 3103

Sequence: 1 DAHKSEVAHRFKDLGEENFK.....TCFAEEGKKLVASQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3103	100.0	609	1	ALBU_HUMAN P02768 homo sapien
2	2942	94.8	600	1	ALBU_MAMMU Q28522 macaca mulie
3	2620	84.6	608	1	ALBU_FELCA P49064 felis silve
4	2562	82.6	608	1	ALBU_CANFA P49822 canis fami
5	2475.5	79.8	607	1	ALBU_HORSE P35747 equus cabai
6	2450.5	79.0	607	1	ALBU_BOVIN P02769 bos taurus
7	2446	78.8	608	1	ALBU_RABIT P49065 oryctolagus
8	2432.5	78.4	607	1	ALBU_SHEEP P14639 ovis aries
9	2426	78.2	608	1	ALBU_RAT P02770 rattus norv
10	2411.5	77.7	605	1	ALBU_PIG P08835 sus scrofa
11	2387	76.9	609	1	ALBU_MERUN O35090 meriones un
12	2378	76.6	609	1	ALBU_MOUSE P07724 mus musculu
13	1557.5	50.2	615	1	ALBU_CHICK P19121 gallus galli
14	1253.5	40.4	609	1	FETA_PANTR Q28759 pan treglod
15	1249.5	40.3	609	1	FETA_HUMAN P02771 homo sapien
16	1242.5	40.0	609	1	FETA_GORGO P28050 gorilla gor
17	1205	38.8	607	1	ALB2_XENLA P14872 xenopus lae
18	1200	38.7	609	1	FETA_HORSE P43066 equus cabai
19	1164.5	37.5	606	1	A-B1_XENLA P08759 xenopus lae
20	1084	34.9	605	1	FETA_MOUSE P02772 mus musculu
21	1067	34.4	611	1	FETA_RAT P02773 rattus norv
22	1055	34.0	599	1	AFAM_HUMAN O43652 homo sapien
23	944	30.4	611	1	AFAM_MOUSE O83020 mus musculu
24	928	29.9	608	1	AFAM_RAT P36953 rattus norv
25	747.5	24.1	608	1	ALB1_SALSA P21848 salmo salar
26	742.5	23.9	608	1	ALB2_SALSA Q03156 salmo salar
27	699	22.5	182	1	ALBU_RANCA P21847 rana catesb
28	440.5	14.2	1423	1	ALBU_PETMA Q91274 petronyzen
29	386	12.4	474	1	VTDB_HUMAN P02774 homo sapien
30	381	12.3	476	1	VTDB_RAT P04276 rattus norv
31	378	12.2	476	1	VTDB_RABIT P53789 oryctolagus
32	372	12.0	472	1	VTDB_MOUSE P21614 mus musculu
33	151.5	4.9	1605	1	Q99p15 mus musculu

34	144.5	4.7	8797	1	SNE1_HUMAN Q8nf91 homo sapien
35	138.5	4.5	1410	1	RRB1_HUMAN Q9p282 homo sapien
36	133.5	4.3	1391	1	NST3_DROHY Q08696 drosophila
37	132.5	4.3	2230	1	GOG4_HUMAN Q13435 homo sapien
38	129.5	4.2	1972	1	MYHB_HUMAN P15749 homo sapien
39	129	4.2	3210	1	CENF_HUMAN P49454 homo sapien
40	128	4.1	1005	1	RA50_METJA Q58718 methanococc
41	126.5	4.1	1972	1	MYHB_RABIT P15748 oryctolagus
42	126	4.1	1189	1	SMC2_CHICK Q90988 galus galli
43	126	4.1	3259	1	GIAK_HUMAN Q14789 homo sapien
44	125	4.0	3038	1	TRIO_HUMAN O75962 homo sapien
45	124.5	4.0	1790	1	USO1_YEAST P25386 saccharomyc

ALIGNMENTS

RESULT 1	ALBU_HUMAN	STANDARD:	PRT:	609 AA
AC	P02768; Q95574; C13140; Q9P157; Q9P117; Q9UHS3; Q9JUZC;			
DT	21-JUL-1986 (Rel. 51, Created)			
D7	01-APR-1990 (Rel. 14, Last sequence update)			
DE	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Serum albumin precursor.			
GN	ALB			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86296112; PubMed=3009475;			
RA	Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,			
RA	Beattie W.G., Dugaiczky A.;			
RT	"Molecular structure of the human albumin gene is revealed by			
RT	nucleotide sequence within g11-22 of chromosome 4.";			
RL	J. Biol. Chem. 261:6747-6757(1986);			
RN	[2]			
RP	SEQUENCE FROM N.A., AND VARIANT LYS-420.			
RX	MEDLINE=82081882; PubMed=6171778;			
RA	Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,			
RA	Najarian R.C., Seeburg P.H., Wion K.L.;			
RT	"The sequence of human serum albumin cDNA and its expression in E.			
RT	coli.";			
RL	Nucleic Acids Res. 9:6103-6114(1981);			
RN	[3]			
RP	SEQUENCE FROM N.A., AND VARIANT GLY-121.			
RX	MEDLINE=82105994; PubMed=6275391;			
RA	Dugaiczky A., Law S.W., Dennison O.E.;			
RT	"Nucleotide sequence and the encoded amino acids of human serum			
RT	albumin mRNA.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982);			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RA	Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;			
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A. (PROC903/PRO1708/PRO2044/PRO2619/PRO2675).			
RC	TISSUE=Fetal liver;			
RA	Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,			
RA	Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;			
RT	"Functional prediction of the coding sequences of 121 new genes			
RT	deduced by analysis of cDNA clones from human fetal liver.";			
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.			
RA	Huang M.C., Wu H.T.;			
RT	"The cDNA sequences of human serum albumin.";			
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.			
RN	[7]			
RP	SEQUENCE FROM N.A.			

RC TISSUE=Liver, and Skeletal muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Feingold E.A., Grouse J.H., Derge J.G.,
RA Altshuler S.F., Zellberg B., Suetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S., Kang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollany S.J.,
RA Bosca S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.S., Lu X., Gibbs R.A.,
RA Faney J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [18]
RP SEQUENCE OF 25-609.
RX MEDLINE=76187907; PubMed=1225573;
RA Meloun B., Moravsek L., Kostka V.;
RT "Complete amino acid sequence of human serum albumin";
RL FEBS Lett. 58:134-137(1975).
RN [19]
RP SEQUENCE OF 25-609.
RA Brown J.R., Shockley P., Behrens P.Q.;
RJ (in:) Bing D.H. (eds.);
RL The chemistry and physiology of the human plasma proteins, pp.23-40,
RL Pergamon Press, New York (1979).
RN [10]
RP SEQUENCE OF 1-455 FROM N.A.
RA Menaya J., Parrilla R., Ayuso M.S.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [11]
RP SEQUENCE OF 1-26 FROM N.A.
RX MEDLINE=86:40099; PubMed=2419329;
RA Urano Y., Watanabe K., Sakai T., Tamaoki T.;
RT "The human albumin gene. Characterization of the 5' and 3' flanking
regions and the polyorphic gene transcripts";
RL J. Biol. Chem. 261:3244-3251(1986).
RN [12]
RP SEQUENCE OF 222-229.
RX MEDLINE=76257808; PubMed=955075;
RA Walker J.E.;
RT "Lysine residue 199 of human serum albumin is modified by
acetylsalicylic acid";
RL FEBS Lett. 66:173-175(1976).
RN [13]
RP SEQUENCE OF 25-44 AND 48-499.
RX MEDLINE=95203287; PubMed=7895732;
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT "The human myocardial two-dimensional gel protein database: update
1994";
RL Electrophoresis 15:1459-1465(1994).
RN [14]
RP SEQUENCE OF 25-44 AND 48-499.
RX MEDLINE=95203287; PubMed=7895732;
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT "The human myocardial two-dimensional gel protein database: update
1994";
RL Electrophoresis 15:1459-1465(1994).
RN [15]
RP SEQUENCE OF 25-44 AND 48-499.
RX MEDLINE=78186630; PubMed=656055;
RA Jacobsen C.;
RT "Lysine residue 240 of human serum albumin is involved in high-
affinity binding of bilirubin";
RL Biochem. J. 171:453-459(1978).
RN [16]
RP VARIANT CANTERBURY ASN-337.
RX MEDLINE=87157744; PubMed=3828358;
RA Brennan S.O., Herbert P.;
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
domain of serum albumin";
RL Biochim. Biophys. Acta 912:191-197(1987).
RN [17]
RP VARIANTS NAG-2 AND NAG-3.
RX MEDLINE=8068523; PubMed=3479777;
RA Takashashi N., Takashashi Y., Isobe T., Putnam F.W., Fujita M.,
RA Satoh C., Neel J.V.;
RT "Amino acid substitutions in inherited albumin variants from
Amerindian and Japanese populations";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
RN [18]
RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.
RX MEDLINE=89345611; PubMed=2762316;
RA Arai K., Madison J., Huss K., Ishioke N., Satch C., Fujita M.,
RA Neel J.V., Sakurabayashi I., Putnam F.W.;
RT "Point substitutions in Japanese alloalbumins";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
RN [19]
RP VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW GUINEA.
RX MEDLINE=90115905; PubMed=2404284;
RA Arai K., Madison J., Shimizu A., Putnam F.W.;
RT "Point substitutions in albumin genetic variants from Asia";
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
RN [20]
RP DESCRIPTION OF VARIANT REDHILL.
RX MEDLINE=90115852; PubMed=2104980;
RA Brennan S.O., Myles T., Peach R.J., Donaldson J., George P.M.;
RT "Albumin Redhill (1 Arg, 320 Ala-->Thr): a glycoprotein variant of
human serum albumin whose precursor has an aberrant signal peptidase
cleavage site";
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
RN [21]
RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
RX MEDLINE=91062352; PubMed=2247440;
RA Galliano M., Michiotti L., Porta F., Rossi A., Ferri G., Madison J.,
RA Watkins S., Putnam F.W.;
RT "Mutations in genetic variants of human serum albumin found in
Italy";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
RN [22]
RP VARIANT VENEZIA.
RX MEDLINE=91296740; PubMed=2068071;
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
RA Minchiotti L., Putnam F.W.;
RT "A donor splice mutation and a single-base deletion produce two
carboxyl-terminal variants of human serum albumin";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
RN [23]
RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;
RX KOMAGOME-2 ARG-152 AND KOMAGOME-1 GIU-396.
RX MEDLINE=92052189; PubMed=1946412;
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
RA Matsuda Y., Amaki I., Putnam F.W.;
RT "Genetic variants of serum albumin in Americans and Japanese";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
RN [24]
RP VARIANT CASEBROOK ASN-518.
RX MEDLINE=91316157; PubMed=1859851;
RA Peach R.J., Brennan S.O.;
RT "Structural characterization of a glycoprotein variant of human serum
albumin: albumin Casebrook (494 Asp-->Asn)";
RL Biochim. Biophys. Acta 1097:49-54(1991).
RN [25]
RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
RX MEDLINE=92190239; PubMed=1347703;
RA Mirchioti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
RA Rochu D., Porta F.;
RT "Two alloalbumins with identical electrophoretic mobility are produced


```

Query Match      100.0%; Score 3103; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.7e-194;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFDLGBENFKALVLIAPAOYLQOCPEFHVKLVNVTFFAKTCVADESSE 60
DB 25 DAHKSVAHRFDLGBENFKALVLIAPAOYLQOCPEFHVKLVNVTFFAKTCVADESSE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPCLPRVPR 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPCLPRVPR 144
QY 121 DVNCTAPHONEETFLKKYLVEARRHPYFAPPELLFFAKRYKAATFECQAAKACLLP 180
DB 145 DVNCTAPHONEETFLKKYLVEARRHPYFAPPELLFFAKRYKAATFECQAAKACLLP 204
QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQFPKAEFAEVSCLVDTLTK 240
DB 205 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQFPKAEFAEVSCLVDTLTK 264
QY 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVSKDCVCKNVAEAKDVLGMLFVEYARRHPDYSVWLLRLAKTYETLLEKC 360
DB 325 DLPSLAADFVSKDCVCKNVAEAKDVLGMLFVEYARRHPDYSVWLLRLAKTYETLLEKC 384
QY 361 CAADAPHECYAKVDFEFLKPLVEEPQNLKQNCLEFQGLGEYKFNQALLVRYTKVPQYST 420
DB 385 CAADAPHECYAKVDFEFLKPLVEEPQNLKQNCLEFQGLGEYKFNQALLVRYTKVPQYST 444
QY 421 PTLVEVSRNLGKVGSKCKRHPKAKRMPCHAEEDVLSVNLGJCVLHREKTPVSDRYTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKRHPKAKRMPCHAEEDVLSVNLGJCVLHREKTPVSDRYTKCCTES 504
QY 481 LVNRPCFSALEVDYTVPKFNAETFTPHADICTLSEKERIKKOTALVELVHKPKAT 540
DB 505 LVNRPCFSALEVDYTVPKFNAETFTPHADICTLSEKERIKKOTALVELVHKPKAT 564
QY 541 KEOLKAWMDPFAAFVKKCKADDKETCFABEGKLVAAQAALGL 585
DB 565 KEOLKAWMDPFAAFVKKCKADDKETCFABEGKLVAAQAALGL 609

RESULT 2
ALBU_MACMU STANDARD; PRF; 600 AA.
AC Q28522;
AT 01-NOV-1997 (Rel. 35, Created);
DT 01-NOV-1997 (Rel. 35, Last sequence update);
DT 28-FEB-2003 (Rel. 41, Last annotation update);
DE Serum albumin precursor (Fragment);
GN A.B.
OS Macaca mulatta (Rhesus macaque);
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
CC Cercopithecinae; Macaca.
CX NCBI_TaxID:9544;
RN [1]_TaxID:9544;
RP SEQUENCE FROM N.A.
RX MEDLINE:93211971; PubMed:8460152;
RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
RA Dwalet J., Putnam F.W.;
RC "cDNA and protein sequence of polymorphic macaque albumins that differ
RT in bilirubin binding.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).
CC -!- FUNCTION: Serum albumin; the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloid osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.

```

```

-!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: V90463; AAA36906.1; -
CC PR: A47391; A47391.
CC HSP: P02769; IE7B.
CC InterPro: IPR000264; Serum albumin.
CC Pfam: PF00273; Transport Prot. 3.
CC PRINTS: PR00802; SERUMALBUMIN.
CC ProDom: PDC02486; Serum albumin; 1.
CC SMART: SM00103; ALBUMIN_3.
CC PROSITE: PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 10 BY SIMILARITY.
FT PROPEP 11 16 BY SIMILARITY.
FT CHAIN 17 600 SERUM ALBUMIN.
FT DOMAIN 17 197 ALBUMIN 1.
FT DOMAIN 204 389 ALBUMIN 2.
FT DOMAIN 396 587 ALBUMIN 3.
FT METAL 19 19 COPPER (BY SIMILARITY);
FT BINDING 256 256 BILIRUBIN (POTENTIAL);
FT DISULFID 69 78 BY SIMILARITY.
FT DISULFID 91 107 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 140 185 BY SIMILARITY.
FT DISULFID 194 193 BY SIMILARITY.
FT DISULFID 216 262 BY SIMILARITY.
FT DISULFID 261 269 BY SIMILARITY.
FT DISULFID 281 295 BY SIMILARITY.
FT DISULFID 294 305 BY SIMILARITY.
FT DISULFID 332 377 BY SIMILARITY.
FT DISULFID 376 385 BY SIMILARITY.
FT DISULFID 408 454 BY SIMILARITY.
FT DISULFID 453 464 BY SIMILARITY.
FT DISULFID 477 493 BY SIMILARITY.
FT DISULFID 492 503 BY SIMILARITY.
FT DISULFID 530 575 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
SQ SEQUENCE 600 AA; 67880 MW; 545C871A670E740B CRC64;

Query Match      94.8%; Score 2942; DB 1; Length 600;
Best Local Similarity 93.5%; Pred. No. 4.8e-184;
Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFDLGBENFKALVLIAPAOYLQOCPEFHVKLVNVTFFAKTCVADESSE 60
DB 17 DTHKSEVAHRFDLGBENFKALVLIAPAOYLQOCPEFHVKLVNVTFFAKTCVADESSE 76
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPCLPRVPR 120
DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPCLPRVPR 136
QY 121 DVNCTAPHONEETFLKKYLVEARRHPYFAPPELLFFAKRYKAATFECQAAKACLLP 180
DB 137 DVNCTAPHONEETFLKKYLVEARRHPYFAPPELLFFAKRYKAATFECQAAKACLLP 196
QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQFPKAEFAEVSCLVDTLTK 240
DB 197 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQFPKAEFAEVSCLVDTLTK 256
QY 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 257 VHTTECHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 316

```


QY 541 KEOLKAVMDPFAAFVEKCKADKCTCFABEGKLVAAQAAL 583
 D 565 DEQLKTVMDFGNFVEKCKAENKEGCFSEBGPFLVNAQAAL 607

RESULT 5

ALBU_HORSE STANDARD; PRT; 607 AA.
 AC P35737;
 DT 01-JUN-1994 (Rel. 23, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor.
 GN ALB.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
 RP TISSUE=Liver;
 RX MEDLINE=93345495; PubMed=8344282;
 RA Ho J.A., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;
 RT "X-ray and primary structure of horse serum albumin (Equus caballus)
 at 0.27-nm resolution."
 RL Eur. J. Biochem. 215:205-212(1993).
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 hormones, bilirubin and drugs. Its main function is the regulation
 of the colloidal osmotic pressure of blood.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Plasma.
 CC -1- SIMILARITY: BELONGS TO THE ALB/AFP/VDS FAMILY.
 CC -1- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL: X74045; CAA52194.1; -
 DR PIR: S34053; ABHOS.
 DR HSP: P02768; 157B.
 DR InterPro: IPR000264; Serum albumin.
 DR Pfam: PF00273; transport_prot_3.
 DR PRINTS: PR00902; SERUMALBUMIN.
 DR ProDom: PD002496; Serum albumin; 1.
 DR SMART: SM00103; ALBUMIN; 3.
 DR PROSITE: PS00212; ALBUMIN; 3.
 KW Meta-binding; Lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 607
 FT DOMAIN 25 204
 FT DOMAIN 211 396
 FT DOMAIN 403 594
 FT META 27 27
 FT META 27 27
 FT DISULFID 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 147 192
 FT DISULFID 191 200
 FT DISULFID 223 269
 FT DISULFID 268 276
 FT DISULFID 288 302
 FT DISULFID 301 312
 FT DISULFID 339 384
 FT DISULFID 383 392
 FT DISULFID 415 461

FT DISULFID 460 471
 FT DISULFID 484 500
 FT DISULFID 499 510
 FT DISULFID 537 582
 FT DISULFID 581 590
 SQ SEQUENCE 607 AA; 68598 MW; 256F6E810A1B90C5 CRC64;
 Query Match: 79.8%; Score 2475.5; DB 1; Length 607;
 Best Local Similarity 76.3%; Pred. No. 9.8e-154;
 Matches 445; Conservative 69; Mismatches 68; Indels 1; Gaps 1;
 QY 1 DAHXSEVAHRFKOLGSENFALVIAFAVQLQCPPEFHVKLVNEVTEFAKTCVADESAS 60
 DB 25 DTHXSEIAHRFNDLGEKHFGLVAFPSQYLQCPPEFHVKLVNEVTEFAKTCVADESAS 84
 QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFQLQKDDNENLDELVRPEV 120
 DB 85 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFQLQKDDNENLDELVRPEV 143
 QY 121 DVMCTAFHDMETFLKKLYLZARRHPHYFYAPJELFFAKRYKAAFTCCQAADKAACLLP 180
 DB 144 DAQCAAFQEDPDKFLGKYLVEVAERHPHYFYGPPELLFHABEYKADFTCCPADDKLAELIP 203
 QY 181 KLDLREDEGKASSAKQRLKCSLQKQGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
 DB 204 KLDALKERILLSSAKERLKCSPQNGERAVKAWAVARLSQRPKAEFAEVSKLVTDLTK 263
 QY 241 VHTCCCHGDLLECADRADLAKYICENODSISSKLKCECKPILKSHC-AEVENDEMPA 300
 DB 264 VHKCCCHGDLLECADRADLAKYICEHODSISGLKACCDKPLQKSHCIAEVKEDDLP 323
 QY 301 DLPSLAADFVESKDVCKYAEAKDVFLGMFLYELAYARRHPDYVSVVLLRLAKTYETLEK 360
 DB 324 DLPALAADFADKEICKHYKDAKDVFLGTFLYELAYARRHPDYVSVVLLRLAKTYETLEK 383
 QY 361 CAADPHCEYAKVDFEPKPLVEEPQNLKQNCFLFEOLGEYKFNALVRYTKVPOVST 420
 DB 394 CAEADPPACVTVDFDQFTPLVEEPKSLVKNKNCDFEEVGYDFONALIVRYTKAPOVST 443
 QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCADYLSVWLNQLCVLREKTPVSDRVTKCCTES 480
 DB 444 PTLVEIGTGLKVGSRCKLPESERLPCSENHLALNRLCVLHEKTPVSEKITKCTDS 503
 QY 481 LVNRRPCFSALEVBETVVPKFEFNAETTFHADICTLSEKERQIKKQALVELVKKPKAT 540
 DB 504 LAERRPCFSALELDEGVVPKFEKAEFTTFHADICTLPEDEKQIKKQALVELVKKPKAT 563
 QY 541 KEOLKAVMDPFAAFVEKCKADKCTCFABEGKLVAAQAAL 583
 DB 564 DEQLKTVMDFGNFVEKCKAENKEGCFSEBGPFLVNAQAAL 606

RESULT 6

ALBU_BOVIN STANDARD; PRT; 607 AA.
 ID ALBU_BOVIN STANDARD; PRT; 607 AA.
 AC P02769; O02787;
 DT 21-JUL-1986 (Rel. 31, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor (Allergen Bos d 6).
 GN ALB.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]_TaxID=9913;
 RP SEQUENCE FROM N.A.
 RA Holowachuk E.W., Stoltenborg J.K., Reed R.G., Peters T. Jr.;
 RL Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANT THR-214.
 RC TISSUE=Liver;

RA Barry T., Power S., Gannon F.;
 R1 Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
 R2 [3]
 R3 SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A., AND VARIANT THR-214.
 RA Wu H.T., Huang M.C.;
 R1 "The complete cDNA sequence of bovine serum albumin."
 R2 Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 1-32.
 RX MEDLINE=80024273; PubMed=488109;
 RA McGillivray R.T.A., Chung D.W., Davies E.W.;
 R1 "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-
 R2 terminal sequence of preproalbumin."
 RL Eur. J. Biochem. 98:477-485(1979);
 RN [6]
 RP SEQUENCE OF 25-424 AND 423-607, AND VARIANT THR-214.
 RA Brown J.R.;
 R1 "Structure of bovine serum albumin."
 R2 Fed. Proc. 34:591-591(1975).
 RN [7]
 RP REVISIONS TO 190-195.
 RA Brown J.R.;
 R1 Submitted (APR-1975) to the PIR data bank.
 RN [8]
 RP SEQUENCE OF 402-433.
 RX MEDLINE=82023364; PubMed=7283978;
 RA Reed R.G., Putnam F.W., Peters T. Jr.;
 R1 "Sequence of residues 403-403 of bovine serum albumin."
 RL Biochem. J. 191:867-868(1980).
 RN [9]
 RP SEQUENCE OF 19-28.
 RX MEDLINE=77134075; PubMed=849354;
 RA Patterson J.E., Geiler D.M.;
 R1 "Bovine microsomal albumin: amino terminal sequence of bovine
 R2 proalbumin."
 RL Biochem. Biophys. Res. Commun. 74:1220-1226(1977).
 RN [10]
 RP SEQUENCE, AND REVISIONS TO 118-119 AND 180.
 RX MEDLINE=91083649; PubMed=2260975;
 RA Hirayama K., Akashi S., Furuya M., Fukuhara K.-I.;
 R1 "Rapid confirmation and revision of the primary structure of bovine
 R2 serum albumin by EIMS and Frit-FAB LC/MS."
 RL Biochem. Biophys. Res. Commun. 173:639-646(1990).
 RN [11]
 RP SEQUENCE OF 25-41.
 RX MEDLINE=88267456; PubMed=3389500;
 RA Hsieh J.C., Lin F.P., Tam M.F.;
 R1 "Electroblotting onto glass-fiber filter from an analytical
 R2 isoelectrofocusing gel: a preparative method for isolating proteins
 R3 for N-terminal microsequencing."
 RL Anal. Biochem. 170:11-8(1988).
 RN [12]
 RP SEQUENCE OF 437-451.
 RA Vilbois F.;
 R1 Submitted (AUG-1998) to the SWISS-PROT data bank.
 RN [13]
 RP DISULFIDE BONDS.
 RA Brown J.R.;
 R1 "Structure of serum albumin: disulfide bridges."
 RL Fed. Proc. 33:1389-1389(1974).
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloid osmotic pressure of blood.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Plasma.
 CC -1- SIMILARITY: BELONGS TO THE ALB/AF/VDL FAMILY.
 CC -1- SIMILARITY: Contains 3 albumin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X73993; AAA51411.1; -;
 DR EMBL; X58989; CAA41735.1; -;
 DR EMBL; Y17769; CAA76847.1; -;
 DR EMBL; AF542068; AAM17824.1; -;
 DR HSSP; P02768; 1E7B;
 DR InterPro; IPR000264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR ProTis; PRO0802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 DR Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;
 KW Peptidomorphism.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 607 SERUM ALBUMIN.
 FT DOMAIN 1 25 ALBUMIN 1.
 FT DOMAIN 2 21 396 ALBUMIN 2.
 FT DOMAIN 3 403 594 ALBUMIN 3.
 FT METAL 27 27 COPPER (BY SIMILARITY).
 FT DISULFID 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 147 192
 FT DISULFID 191 200
 FT DISULFID 223 269
 FT DISULFID 268 276
 FT DISULFID 288 302
 FT DISULFID 301 312
 FT DISULFID 339 384
 FT DISULFID 383 392
 FT DISULFID 415 461
 FT DISULFID 460 471
 FT DISULFID 484 500
 FT DISULFID 499 510
 FT DISULFID 537 582
 FT DISULFID 581 590
 FT VARIANT 214 214
 FT CONFLICT 302 302 A -> T.
 FT CONFLICT 304 305 C -> K (IN REF. 6).
 FT CONFLICT 324 324 KP -> PC (IN REF. 6).
 FT CONFLICT 324 324 N -> D (IN REF. 6).
 FT CONFLICT 394 395 ST -> TS (IN REF. 6).
 FT CONFLICT 437 437 K -> R (IN REF. 12).
 FT CONFLICT 493 494 SE -> ES (IN REF. 6).
 SQ SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;
 Query Match 79.0%; Score 2450.5; DB 1; Length 607;
 Best Local Similarity 75.8%; Pred. No. 41e-152;
 Matches 442; Conservative 71; Mismatches 69; Indels 1; Gaps 2;
 QY 1 DAHKSVAHFRFKDLGSENFKAJVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESSE 60
 DB 25 DTHKSEIAHFRFKDLGSEHFKGLVLIAPFQYLOQCPEDHVKLVNELTEFAKTCVADESHA 84
 QY 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 DB 85 GCEKSLHTLFGDELCKVASRETYGEMADCCCKQEPERNECFLOHSHKDDSPDLPLKLP 143
 QY 121 DVMCTAFTHDEHTEFLKKLYLVEIARRHPYFYAPLLEJFFAKYKAAFTCCQAAADKACLDP 180
 DB 144 NTLCDPEKADKKFKWGLYELIARRHPYFYAPLLEJYANKYNGVFQCCQADKGLCLLP 203
 QY 161 KLDELDEGVASSAKQKLKCLASLOKGERAFKAWAVARLSORFPKAEFAEVSKLVTLDTLK 240

Db 204 KIETREKVLASSARQCRJRCASIQKFGERALKAWSVARLSQKFPKAEFVEVTKLVDTLTK 263

QY 241 VHTCECHGDLLECCADBRADLAKYICENQDSISSKLEKCECKPILLESKSHCIAEVENDEMPA 300

Db 264 VHKCECHGDLLECCADBRADLAKYICENQDSISSKLEKCECKPILLESKSHCIAEVEKDAPE 323

QY 301 DPLSLAARFVESKVCNKYAAEKDFLGMFLVEYARRHPDYSVVLRLRIAKTYETTLK 360

Db 324 NLPLTADPAEDKVCNKYQENKDAFLGSFLVEYSRHPDYAVSVLRLAKYEATLEK 383

QY 361 CAADPHCEYAKVDFDKPVEBPQNDLQKNCBELFQJAEYKFNALJYVYTKKVPQVST 420

Db 384 CAKDQPHACYSTVDFKLKLVDEPNLQKNCDFEKLGEYFQNALJYVYTRKVPQVST 443

QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPKADYLSVVLNGLCVLHKTYPYSDRVTKCCTES 480

Db 444 PTLVEVSRNLGKVGSKCKHPEAKRMPKADYLSVVLNGLCVLHKTYPYSDRVTKCCTES 503

QY 481 LVNRRPCFSALEVDETYVYPKFNATFTFHADICTLSEKERQTKQTALVELVGHKPKAT 540

Db 504 LVNRRPCFSALEVDETYVYPKFNATFTFHADICTLSEKERQTKQTALVELVGHKPKAT 563

QY 541 KEQLKAVMDDFAFVPEKCKKADDKETCFABEGKLUVAASQAL 593

Db 564 EQQLKTMENFAVFDKCAADDKETCFABEGKLUVSTQAL 606

RESULT 7

ALBU_RABIT STANDARD; PRT; 608 AA.

AC P49065;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Serum albumin precursor.

GN ALB.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=New Zealand white; TISSUE=Liver;

RA Subfield W.P., Syed S., Schuyler P.C.;

RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good

CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,

CC hormones, bilirubin and drugs. Its main function is the regulation

CC of the colloidal osmotic pressure of blood.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Plasma.

CC -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.

CC -!- SIMILARITY: Contains 3 albumin domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

CC EMBL; U18344; AAB58347.1; -

DR HSSP; P02768; iE75.

DR InterPro: IPR000264; Serum albumin.

DR Pfam: PF00273; transport_prot; 3.

DR PRINTS; PR00802; SERJALBUMIN.

DR ProDom; PD002486; Serum albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.

KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.

FT SIGNAL 1 18 BY SIMILARITY.

FT PROPEP 19 24 BY SIMILARITY.

FT CHAIN 25 608 SERUM ALBUMIN.

FT DOKAIN 25 265 ALBUMIN 1.

FT DOKAIN 212 397 ALBUMIN 2.

FT DOKAIN 404 595 ALBUMIN 3.

FT METAL 27 27 COPPER.

FT DISULFID 77 86 BY SIMILARITY.

FT DISULFID 99 115 BY SIMILARITY.

FT DISULFID 114 125 BY SIMILARITY.

FT DISULFID 148 193 BY SIMILARITY.

FT DISULFID 192 201 BY SIMILARITY.

FT DISULFID 224 270 BY SIMILARITY.

FT DISULFID 269 277 BY SIMILARITY.

FT DISULFID 289 303 BY SIMILARITY.

FT DISULFID 302 313 BY SIMILARITY.

FT DISULFID 340 385 BY SIMILARITY.

FT DISULFID 384 393 BY SIMILARITY.

FT DISULFID 416 462 BY SIMILARITY.

FT DISULFID 461 472 BY SIMILARITY.

FT DISULFID 485 501 BY SIMILARITY.

FT DISULFID 500 511 BY SIMILARITY.

FT DISULFID 538 583 BY SIMILARITY.

FT DISULFID 582 591 BY SIMILARITY.

SQ SEQUENCE 608 AA; CF592647AAFE9A2 CRC64;

Query Match 78.8%; Score 2446; DB 1; Length 608;

Best Local Similarity 74.1%; Pred. No. 8,1e-152;

Matches 433; Conservative 77; Mismatches 74; Indels 0; Gaps 3;

QY 1 DAHSEVARRFKDLGEENFKALVJLAFAYLOQCFEDHVKLVNEVTEFAKTCVADSEAE 60

DB 25 EAHKSEIARFNDVGBEHHFGLVLTITFSYLOKCPYEBHAKLVKEVTDLAKACVADSEAA 84

QY 61 NCDKSLHCTLFQDKLCTVATLRETYGEMADCCAKPEPERNECFLOHKDNDNPLRLVRPEV 120

DB 85 NCDKSLHDIQDKICALPSLRDTYGVADCCKEKEPERNECFLLHKKDKDPOLPFAPEA 144

QY 121 DVMCTAFHNDNEETFLKLYLYEIAIRRHYPFYAPPELLFFAKYKAAFTCCQADKACLLP 180

DB 145 DVLCKAFHDEKAFPHLYVEVARRHPFYAPPELLLYAOKYKAILTECCCAAOKGACLT 204

QY 181 KLDELRLDEGKASSAKQRLKCLASLOKFGERAFKAVARLQRPFKAFABEVSCKVTDLT 240

DB 205 KLDALGKSLISAQOERLURCASIOKFGDRAYKAWALVRLSORFPKADFTDISKIVTDLT 264

QY 241 VHTCECHGDLLECCADBRADLAKYICENQDSISSKLEKCECKPILLESKSHCIAEVENDEMPA 300

DB 265 VHKCECHGDLLECCADBRADLAKYICENQDSISSKLEKCECKPILLESKSHCIAEVEKDAPE 324

QY 301 DPLSLAARFVESKVCNKYAAEKDFLGMFLVEYARRHPDYSVVLRLRIAKTYETTLK 360

DB 325 GJPAVAEEFVEDKQVCKNYEBAKDFLGKFLVEYSRHPDYSVVLRLRIAKTYETTLK 384

QY 361 CAADPHCEYAKVDFDKPVEBPQNDLQKNCBELFQJAEYKFNALJYVYTRKVPQVST 420

DB 385 CATDDPHACYAKULDEFOPLVDEPNLVKNCCELYEQJGVNFQNALVRYTKVPOVST 444

QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPKADYLSVVLNGLCVLHKTYPYSDRVTKCCTES 480

DB 445 PTLVEVSRNLGKVGSKCKHPEAKRMPKADYLSVVLNGLCVLHKTYPYSDRVTKCCTES 504

QY 481 LVNRRPCFSALEVDETYVYPKFNATFTFHADICTLSEKERQTKQTALVELVGHKPKAT 540

DB 505 LSNRRPCFSALEVDETYVYPKFNATFTFHADICTLSEKERQTKQTALVELVGHKPKAT 564

QY 541 KEQLKAVMDDFAFVPEKCKKADDKETCFABEGKLUVAASQALG 584

DB 565 NDQLKTVVGEFTALDDKCCSAEDKEACFAVEGPKLVESKATLG 608

RESULT 8

ALBU SHEEP

ID ALBU SHEEP STANDARD; PRT; 607 AA.

AC P14639;

DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor.
 GN ALB.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=90098888; PubMed=2602150;
 RA Brown W.M., Dziegielewska K.M., Foreman R.C., Saunders N.R.;
 RT "Nucleotide and deduced amino acid sequence of sheep serum albumin.";
 RL Nucleic Acids Res. 17:10495-10495(1989).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- SIMILARITY: BELONGS TO THE ALB/AFB/VDB FAMILY.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC
 DR EYBL; X17055; CAA34903.1; -;
 DR PIR; S06936; ABSHS.
 DR HSSP; P02768; 1E7E.
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18 BY SIMILARITY.
 FT PROPEP 19 24 BY SIMILARITY.
 FT CHAIN 25 607 SERUM ALBUMIN.
 FT DOMAIN 25 204 ALBUMIN 1.
 FT DOMAIN 211 396 ALBUMIN 2.
 FT DOMAIN 403 594 ALBUMIN 3.
 FT METAL 27 27 COPPER (BY SIMILARITY).
 FT DISULFID 77 86 BY SIMILARITY.
 FT DISULFID 99 115 BY SIMILARITY.
 FT DISULFID 114 125 BY SIMILARITY.
 FT DISULFID 147 192 BY SIMILARITY.
 FT DISULFID 191 200 BY SIMILARITY.
 FT DISULFID 223 269 BY SIMILARITY.
 FT DISULFID 268 276 BY SIMILARITY.
 FT DISULFID 288 302 BY SIMILARITY.
 FT DISULFID 301 312 BY SIMILARITY.
 FT DISULFID 339 384 BY SIMILARITY.
 FT DISULFID 383 392 BY SIMILARITY.
 FT DISULFID 415 461 BY SIMILARITY.
 FT DISULFID 460 471 BY SIMILARITY.
 FT DISULFID 484 500 BY SIMILARITY.
 FT DISULFID 499 510 BY SIMILARITY.
 FT DISULFID 537 582 BY SIMILARITY.
 FT DISULFID 581 590 BY SIMILARITY.
 SQ SEQUENCE 607 AA; 69198 MW; 84979AB7FB86596 CRC64;
 Query Match 78.4%; Score 2432.5; DB 1; Length 607;
 Best Local Similarity 75.0%; Pred. No. 6.1e-151;
 Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

QY 1 DAHKSEVAHRFKDLGEENFKALVLJAFACYLQQCFEDFHVKLNVETFAKTQVADSAAE 60
 DB 25 DTHKSEIAHRFNDLGEENFQGLVLJAFASQYLQQCFDFDEHVKLNVKELTEFAKTQVADSAAE 84
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLPRVREV 120
 DB 85 GCDKSLHTLFGDELCKVATLRETYGDMADCKEKEPERNECFLNHKDNPDLPLK-RPEP 143
 QY 121 DVMCTAFHDNEETFLKKYLYEIAHRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 DB 144 DTLCAEFKADKKFWGKLYEVARRHPYFYAPPELLYYANKNGVFOECCQAEKGCALLP 203
 QY 181 KLDELRLDECKASSAKORLKCAASLOKFGESAPKANAVARLSQRFPAKAEFAEYSKLVTLTK 240
 DB 204 KIDAMREKVLASSARORLCASIQKFGERALKAMSVARLSQKFPKACFTVTK-VTDLTK 263
 QY 241 VHTCECHGDLLECADRADLAKYICENODSISSKLKECCXPLLEKSHCIAEVENDEMPA 300
 DB 264 VHKCECHGDLLESCADRADLAKYICDHCDALUSSKLKECCDKRPVLEKSHCIAEVDKDAVPE 323
 QY 301 DLPLSLAADFVESKDYCKNYAKAKOVFLQMFYEVARRHPDYVSVLLRLAKTYETTLK 360
 DB 324 NLPLPLTADPAEDKVEKYNQAEKOVFLQSGFLYEXSRHPPEYAVSVLLRLAKEYEATLEDC 383
 QY 361 CAAADPHECYAKVDFEKPVEEPQNLKONCELFEOLGEYKFGONALVRVTKKVPQVST 420
 DB 384 CAKEDPHACYATVFDKLKHLVDPEQNLKONCELFEKGEYFGONALVRVTRKAPQVST 443
 QY 421 PTLVEVSRNLGKVGSKCKHPKAPKAPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTTES 480
 DB 444 PTLVEISRLGKVGKCKAPKAPSERMPCTEDYLSILNRLCVLHEKTPVSEKVTCTTES 503
 QY 481 LVNRRPCFSALEVDYTYPKFNAETFTPHAD:CTLSEKERCIKQTALVELVGHKPKAT 540
 DB 504 LVNRRPCFSDLTLDYTYPKFPEKFFTFPHADICTFLPDTXQIKQTALVELVGHKPKAT 563
 QY 541 KEOLKAVMDQDFAAFVEKCKCKADDKETCTFAEEGKCLVAAASQAAL 583
 DB 564 DEOLKTVMENFVAFVDKCCAADKSGCFVLEGPKVAVASTQAAL 606
 RESULT 9
 ALBU RAT
 ID ALBU RAT STANDARD; PRG; 608 AA.
 AC P02770; P11382;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP);
 GN ALB.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81223722; PubMed=7017712;
 RA Sargent T.D., Yang W., Bonner J.;
 RT "Nucleotide sequence of cloned rat serum albumin messenger RNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).
 RN [2]
 RP SEQUENCE OF 1-38, AND PROCESSING.
 RX MEDLINE=7249657; PubMed=893447;
 RA Strauss A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;
 RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-
 RT piece. Analysis of the direct translation product of albumin
 RT messenger RNA.";
 RL J. Biol. Chem. 252:6846-6855(1977).
 RN [3]
 RP SEQUENCE OF 25-222.
 RX MEDLINE=78109429; PubMed=564345;
 RA Isemura S., Ikenaka T.;

RT "Amino acid sequences of fragments I and II obtained by cyanogen
 RL bromide cleavage of rat serum albumin.";
 RN J. Biochem. 83:35-48(1978).
 RP [4]
 RP SEQUENCE OF 223-288 AND 572-608.
 RA MEDLINE=76260153; PubMed=956149;
 RA Ikenaka S., Ikenaka T.;
 RT "Fragmentation of rat serum albumin by cyanogen bromide cleavage and
 RL the amino acid sequences of four fragments.";
 RL J. Biochem. 79:1183-1196(1976).
 RN [5]
 RN SEQUENCE CF 168-174.
 RC TISSUE=Plasma;
 RX MEDLINE=87194805; PubMed=2437117;
 RA Carraway R.E., Mitra S.P., Cochrane D.E.;
 RT "Structure of a biologically active neurotensin-related peptide
 RL obtained from pepsin-treated albumin(s).";
 RL J. Biol. Chem. 262:5968-5973(1987).
 RN [6]
 RN COPPER-BINDING.
 RX MEDLINE=79001617; PubMed=80265;
 RA Aoyagi Y., Ikenaka T., Ichida F.;
 RT "Copper(II)-binding ability of human alpha-fetoprotein.";
 RL Cancer Res. 38:3483-3486(1978).
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloid osmotic pressure of blood.
 CC -1- FUNCTION: NRP REGULATES FAT DIGESTION, LIPID ABSORPTION, AND
 CC BLOOD FLOW (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Plasma.
 CC -1- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
 CC -1- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 DR EMBL; V01222; CAA24532.1; -;
 DR PIR; A93872; ABRTS.
 DR HSSP; P02768; IE7B.
 DR InterPro; IPSC00264; Serum albumin.
 DR Pfam; PF00273; transport_prot_3.
 DR PRINTS; PR08802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 608
 FT PEPTIDE 166 174
 FT DOMAIN 1 25
 FT DOMAIN 2 212
 FT DOMAIN 3 404
 FT METAL 27 27
 FT DISULFID 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 148 193
 FT DISULFID 192 201
 FT DISULFID 224 270
 FT DISULFID 269 277
 FT DISULFID 289 303
 FT DISULFID 302 313
 FT DISULFID 340 385
 FT DISULFID 384 393
 FT DISULFID 416 462

FT DISULFID 461 472 BY SIMILARITY.
 FT DISULFID 485 501 BY SIMILARITY.
 FT DISULFID 500 511 BY SIMILARITY.
 FT DISULFID 538 583 BY SIMILARITY.
 FT DISULFID 582 591 BY SIMILARITY.
 FT VARIANT 262 262 V -> L.
 FT CONFLICT 174 174 Y -> S (IN REF. 5).
 SQ SEQUENCE 608 AA; 68718 MW; 538497A282411AB7 CRC64;
 Query Match 78.2%; Score 2426; DB 1; Length 608;
 Sest Local Similarity 73.4%; Pred. No. 1.6e-150;
 Matches 428; Conservative 82; Mismatches 73; Indels 0;
 Gaps 0;
 QY 1 DAKSEVAHREFKDLGEENFKALVLIATFAQYQQCFPFDDHVKLVNEVTEFAKTCVADESA 60
 DB 25 BAKSEIAHREFKDLGEGHFKGLVLIATFSQYLQKPFYEEHKLKVGVEITDFAKTCVADENAE 84
 QY 61 NCDKSLHTFQDKCTVATLRETYGEVADCCAKGEPERNECFLOHKDNDNPLPLVRPEV 120
 DB 85 NCDKSIHTFQDKLCAIPKLRDNYGELADCCAKGEPERNECFLOHKDNDNPLPLVRPEA 144
 QY 121 DMCTAFHDNEETFKKYLVEIARRHVFYAPPELLFFAKRYKAFTSCCAADGAACLLP 180
 DB 145 EAMCTSPQENTSFJGHVLDHEVARRHVFYAPPELLFYAEKYNELVTQCTESDKAACLT 204
 QY 191 KIDELRDEGKASSAKQRLKCSAQKQGERAFKAWAVARLSQRFPAKAEFAEVSKLVTDLTK 240
 DB 205 KLDVAKERKALVAAVRQRMKCSSMQRFGERAFKAWAVARMSQRFPAEFAEITKLADVT 264
 QY 241 VHTCCGGDLLECADDDRADLAKYICENQDSISSKLKECCXPALLEKSHCIAEVENDEMPA 300
 DB 265 INKECHGDLLECADDDRAELAKYMCENQATISSKLOACCXPVLQKSQCLAEETHDNIPA 324
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEARRHDPDYSVLLRLAKTYETTLK 360
 DB 325 DLPSLAADFVEDKCKNYAEAKDVFGLTFLYESRRHDPDYSVLLRLAKTYETTLK 384
 QY 361 CAADDPHECYAKVDFEKPVEEPQNLKONCEFECLGEYKFNALLVRYTKVPQVST 420
 DB 385 CAEGDPPACYGCVLAEPQPLVEEPKLVKTNKCYELKGEYGFQNAVLRVYTKAPQVST 444
 QY 421 PTLVEVSNLKGVSKCKKHPKAKRMPCAEDYLSVNLQVLHEKTPVSDRVTKCCTES 480
 DB 445 PTLVEAARNLGRVGTCKCTLPEAQRLPCVEDYLSAILNRLCLVHEKTPVSEKVTKCCSGS 504
 QY 481 LVNRRPFCFSALEVDSTVVPKSFNAETFTFHADICTLSEKESQIKKOTALVELVKHXPAT 540
 DB 505 LVERRPFCFSALTVDETYVPKSFNAETFTFHSDICTLPDKKQIKKOTALVELVKHXPAT 564
 QY 541 KEQLKAVMDJFAAFVFEKCKKADDKETCFABEGKLVAAASQAAL 583
 DB 565 EDQLKTVMGDPFAQFVKCKCAADKDNCFATEGPNLVARSKEAL 607
 RESULT 10
 ALEB PIG
 ID ALBU PIG STANDARD; PR7; 605 AA.
 AC P08635; Q23018;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE .Serum albumin precursor (Fragment).
 GN ALB.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver.
 RX MEDLINE=89016582; PubMed=3174440;
 RA Baldwin G.S., Weinstock J.;
 RT "Nucleotide sequence of porcine liver albumin.";

RL Nucleic Acids Res. 16:9045-9045(1989).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; X12422; CAA30970.1; -;
 DR EMBL; M36787; AAB30988.1; -;
 DR PIR; S01382; ABOG.
 DR HSSP; P02768; 1E7H.
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 FT NON_TER 1 1
 FT SIGNAL <1 16 BY SIMILARITY.
 FT PROPEP 17 22 BY SIMILARITY.
 FT CHAIN 23 605 SERUM ALBUMIN.
 FT DOXAIN 23 202 ALBUMIN 1.
 FT DOXAIN 209 394 ALBUMIN 2.
 FT DOXAIN 401 592 ALBUMIN 3.
 FT METAL 31 31 COPPER (BY SIMILARITY).
 FT DISULFD 31 84 BY SIMILARITY.
 FT DISULFD 97 123 BY SIMILARITY.
 FT DISULFD 112 123 BY SIMILARITY.
 FT DISULFD 145 190 BY SIMILARITY.
 FT DISULFD 189 198 BY SIMILARITY.
 FT DISULFD 221 267 BY SIMILARITY.
 FT DISULFD 266 274 BY SIMILARITY.
 FT DISULFD 286 300 BY SIMILARITY.
 FT DISULFD 299 310 BY SIMILARITY.
 FT DISULFD 337 382 BY SIMILARITY.
 FT DISULFD 381 390 BY SIMILARITY.
 FT DISULFD 413 459 BY SIMILARITY.
 FT DISULFD 458 469 BY SIMILARITY.
 FT DISULFD 482 498 BY SIMILARITY.
 FT DISULFD 497 508 BY SIMILARITY.
 FT DISULFD 535 580 BY SIMILARITY.
 FT DISULFD 579 588 BY SIMILARITY.
 FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).
 SQ SEQUENCE 605 AA; 69410 MW; 35556B0DD1A1F4FF CRC64;

Query Match
 Best Local Similarity 77.7%; Score 2411.5; DB 1; Length 605;
 Matches 438; Conservative 67; Mismatches 70; Indels 1; Gaps 1;

QY 1 DAHKEVAHFKDGEENFKALVIAFAQYLCQCFEDHVKLVNTEVFPAKTQVADESAE 60
 DB 23 DTJKSEIAHFRKDLGEQYFKGLVLIATFSOHLQCCPQYBEHVKLVREVFPAKTQVADESAE 92
 QY 61 NCDKSLHTLFGDKLCTVATLSETTGEMADCCAKOEBSNECFLOHKCNENLRLVSPV 120
 DB 83 NCDKSLHTLFGDKLCTVATLSETTGEMADCCAKOEBSNECFLOHKCNENLRLVSPV 141
 QY 121 DVMCTAFHNEETFLKKYLYEIAHRRHPYFYAPPELLFPFAKRYKAAFTCCQAADKAACLLP 180
 DB 142 VALCADFCEDEQKFGKYGKLYEIAHRRHPYFYAPPELLYVYAIYKCVFSECCQAADKAACLLP 201

QY 181 KLDELRDEGKASSAKQRLKCSAQKTFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 240
 DB 202 KIEHLREKVLTSAAQRLKCSAQKTFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 261
 QY 241 VHTCCCHGDLLECADRADLAKYICENQDTSISKLKCECKPPLKSHCIAEVENDEMPA 300
 DB 262 VKCCCHGDLLECADRADLAKYICENQDTSISKLKCECKPPLKSHCIAEVENDEMPA 321
 QY 301 DLPSLAADPFVESKDVCKNYABAKDVLGMLFYEYARRHPDYSVLLRLAKTYETTTLEK 360
 DB 322 DLNPLEHDFVEDKEVCNKYKAEKDVFLGTFLYEYARRHPDYSVLLRLAKTYETTTLEK 381
 QY 361 CAADHPCEYAKYVDFEKPVSEPNLIKONCEFLCELYKYNALLVRYTKKVPQVST 420
 DB 382 CAKEEPPACVATVDFKQFPVDBEPKLIKONCEFLCELYKYNALLVRYTKKVPQVST 441
 QY 421 PTLVEVSRNLGKVGSKCKHPEAKRYPCAEDYLSVVLNOLCVLHKEKTPVSDRVTKCTES 480
 DB 442 PTLVEVARKLGLVGRCCCKRPEERLSCAEDYLSVLNRLCVLHKEKTPVSEKVTCTES 501
 QY 481 LVNRRPCFSALEVDETYVPKFNATPTFHADICTLSEKEROIKKQTALVELVHKPKAT 540
 DB 502 LVNRRPCFSALTPTDITYKPEFVEGTFTFHADICTLSEKEROIKKQTALVELVHKPKAT 561
 QY 541 KEOLKAVNDPFAAFVEKCKKADDKETCFABEGKLV 576
 DB 562 ZEQLRTVJGNFAFVQKCAAPDHEACFVGEKGFV 597
 RESULT 11
 ALBU_MERUN
 CD ALBU_MERUN STANDARD; PRT: 609 AA.
 AC O35090;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor.
 GN ALB.
 OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil);
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
 OC Meriones.
 OX NCBI_TaxID=10047;
 RN 1;
 RP SEQUENCE FROM N.A.
 RC STRAIN=MGS IDR; TISSUE=Liver;
 RX MEDLINE=9811663; PubMed=9455485;
 RA Yoshida K., Seto-Onshima A., Sinochawa H.;
 RT "Sequencing of cDNA encoding serum albumin and its extrahepatic
 RT synthesis in the Mongolian gerbil, Meriones unguiculatus.";
 RL DNA Res. 4:351-354(1997).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; AB006197; BAA21765.1; -;
 DR PIR; JCS838; JCS838.
 DR HSSP; P02768; 1E7B.
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transport_prot; 3.

CC -- TISSUE SPECIFICITY: Plasma.
 CC -- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
 CC -- SIMILARITY: Contains 3 albumin domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ011413; CAA09617.1; --
 CC EMBL: M1611; AAA37190.1; --
 CC EMBL: X13063; CAA31458.1; --
 CC EMBL: AK010025; BAB26653.1; --
 CC EMBL: PIR; A35139; A05139.
 CC EMBL: HSSP; P02768; IEB7.
 CC EMBL: SWISS-2DPAGE; P07724; MOUSE.
 CC EMBL: MGD; MGI:87991; Alb1.
 CC InterPro: IPR000264; Serum albumin.
 CC Pfam: PF00273; transprot_prot; 3.
 CC PRINTS: PR0802; SERUMALBUMIN.
 CC ProDom: PD002486; Serum albumin; 1.
 CC SMART: SM00103; ALBUYN; 3.
 CC PROSITE: PS00212; ALBUMIN; 3.
 CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 608
 FT DOMAIN 25 205
 FT DOMAIN 212 397
 FT DOMAIN 404 595
 FT METAL 27 27
 FT DISULFID 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 148 193
 FT DISULFID 192 201
 FT DISULFID 224 270
 FT DISULFID 269 277
 FT DISULFID 289 303
 FT DISULFID 302 313
 FT DISULFID 340 385
 FT DISULFID 384 393
 FT DISULFID 416 462
 FT DISULFID 461 472
 FT DISULFID 500 511
 FT DISULFID 538 583
 FT DISULFID 582 591
 FT CONFLICT 27 27 H -> D (IN REF. 5).
 FT CONFLICT 33 33 H -> D (IN REF. 5).
 FT CONFLICT 41 41 Q -> I (IN REF. 5).
 SQ SEQUENCE 608 AA; 68692 MW; 292F7C7EED3A61B4 CRC64;

Query Match 76.6%; Score 2378; DB 1; Length 608;
 Best Local Similarity 72.4%; Pred. No. 2.1e-147;
 Matches 422; Conservative 80; Mismatches 91; Indels 0; Gaps 0;
 QY 1 DAHSEVAHFKDGLGKGFALVIAFAVQLQCPEDHVKLVNVEVTEFAKTCVADESAA 60
 Db 25 EAHKSEIAHRYNDLGEHFGKGLVLIASFQYKCSYDEHAKLVQEVTEFAKTCVADESAA 84
 QY 61 NCDKSLHTLFGDKLCTVATLRYTGEMADCCAKQEPERNECFQHKDDNPRLVVRPEV 120
 Db 85 NCDKSLHTLFGDKLCAIPNLRENYGELADCCCTQEPERNECFQHKDDNPRLVVRPEV 144
 QY 121 DVNCTAFHNDNEFLKKYLVEIRRPYFYVAPELLFFAKRYKAAPTECCCAKCAKACILP 160
 Db 145 EAMCTGFKENPTTFXGHLVHEVARRHPYFYVAPELLFYAEQYNEILTQCCAEADKSCSLTP 204
 QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLT 240

Db 205 KLGKVKALVSSVRQMKCSSMOKGERAFKAWAVARLSQTPFNADFAETTKUATDLTK 264
 QY 241 VHTCCGGDLLECADRADLAKYICENQDSTISSKKECCCKPLLEKSHCIAEVENDEWPA 300
 Db 265 VNKCCGGDLLECADRADLAKYICENQDSTISSKKECCCKPLLEKSHCIAEVENDEWPA 324
 QY 301 DLPLAADFVSKDVKYKAEAKDVFLGMFLYEVARRHPDYVSVLLRLAKTYETTLKCK 360
 Db 325 DLPLAADFVSKDVKYKAEAKDVFLGMFLYEVARRHPDYVSVLLRLAKTYETTLKCK 384
 QY 361 CAADAPHECYAKVDFEKLVEBPONLIKONCELFEGLEGYKAFONALVRYTKVPOVST 420
 Db 385 CAENPPACVGTVLAEFQPLVEEPONLVKNCDDYKLGEGYGFONALVRYTKVPOVST 444
 QY 421 PTLVEVSRNLGKSGKCKHPEAKMPCAEADYLSVWLNQLCVLHETPVSDRVTKCCTES 480
 Db 445 PTLVEVSRNLGKSGKCKHPEAKMPCAEADYLSVWLNQLCVLHETPVSDRVTKCCTES 504
 QY 481 LVNRRPCFSALEVDVETVVPKEFNAETTFHADICTLSEKEROIKKQALVELVXHKPKAT 540
 Db 505 LVNRRPCFSALEVDVETVVPKEFNAETTFHADICTLSEKEROIKKQALVELVXHKPKAT 564
 QY 541 KEQLKAVMDQFAAFVEKCCCKADDKETCFABEGKGLVAASAAAL 583
 Db 565 AEQLKAVMDQFAAFVEKCCCKADDKETCFABEGKGLVAASAAAL 607

RESULT 13
 ALBU CHICK STANDARD; ERT; 615 AA.
 AC P19:21;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor.
 GN ALB.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI TaxID=9031;
 RN [1] _taxid=9031;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Cassidy A., Salkild C.K., Baverstock P., Wallace J.C.;
 RJ Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE=83161037; PubMed=6187737;
 RA Hache R.J.G., Wiskocil R., Vasa M., Roy R.N., Lau P.C.K., Deeley R.G.;
 RT "The 5' noncoding and flanking regions of the avian very low density
 RT apolipoprotein II and serum albumin genes. Homologies with the egg
 RT white protein genes.";
 RL J. Biol. Chem. 258:4556-4564(1983).
 RN [3]
 RP SEQUENCE OF 19-30.
 RX MEDLINE=78019943; PubMed=911327;
 RA Rosen A.M., Geller D.M.;
 RT "Chicken microsomal albumin: amino terminal sequence of chicken
 RT proalbumin.";
 RL Biochem. Biophys. Res. Commun. 78:1060-1066(1977).
 CC -- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -- SUBCELLULAR LOCATION: Secreted.
 CC -- TISSUE SPECIFICITY: Plasma.
 CC -- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
 CC -- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

DR EMBL: X60688; CAA43098.1; -
 DR EMBL: V00381; CAA23680.1; -
 DR PIR: S15571; ABCSH;
 DR HSP: P02768; IE7B;
 DR InterPro: IPR000264; Serum_albumin.
 DR Pfam: PF00273; transport_prot; 3.
 DR PRINTS: PR00802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum_albumin; 1.
 DR SMART: SM00103; ALBUMIN; 3.
 DR PROSITE: PS00212; ALBUMIN; 3.
 KW Meta-binding; Lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18
 FT PROPEP 19 23
 FT CHAIN 24 615 SERUM ALBUMIN.
 FT DOMAIN 24 209 ALBUMIN 1.
 FT DOMAIN 216 401 ALBUMIN 2.
 FT DOMAIN 408 599 ALBUMIN 3.
 FT METAL 30 30 COPPER (BY SIMILARITY).
 FT DISULFID 80 89 BY SIMILARITY.
 FT DISULFID 102 118 BY SIMILARITY.
 FT DISULFID 117 128 BY SIMILARITY.
 FT DISULFID 152 197 BY SIMILARITY.
 FT DISULFID 196 205 BY SIMILARITY.
 FT DISULFID 228 274 BY SIMILARITY.
 FT DISULFID 273 281 BY SIMILARITY.
 FT DISULFID 293 307 BY SIMILARITY.
 FT DISULFID 306 317 BY SIMILARITY.
 FT DISULFID 344 389 BY SIMILARITY.
 FT DISULFID 386 397 BY SIMILARITY.
 FT DISULFID 420 466 BY SIMILARITY.
 FT DISULFID 465 476 BY SIMILARITY.
 FT DISULFID 489 505 BY SIMILARITY.
 FT DISULFID 504 515 BY SIMILARITY.
 FT DISULFID 542 587 BY SIMILARITY.
 FT DISULFID 586 595 BY SIMILARITY.
 FT CARBOHYD 500 500 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLUENT 24 24 F -> M (IN REF. 3).
 SQ SEQUENCE 615 AA; 69918 MW; E59E4BBAEC066C6 CRC64;

Query Match 50.28; Score 1557.5; DB 1; Length 615;
 Best Local Similarity 46.74; Pred No. 4.3e-94;
 Matches 273; Conservative 118; Mismatches 192; Indels 1; Gaps 1;

```

Qy 3 HKSEVAHRFDLGENFKALVLIIFAQYLOCCPPEDCHVKLVNVEYTFEPAKTCVADESAENC 62
Db 30 HKSEIARVNDLKEETKAVAMITFAQYLQRCVEGLSKLVKQWDLAQKCVANEDAPEK 89
Qy 63 DKSHTLFGDKLCTVATLRETYGEMACCAKQPERNECFCHKDDPNLPR-LVREYD 122
Db 90 SKPLPSILIOEIQCVEXLRDSYGAVACCSKADPERNECFLSFKVSPQDFVQYQFPA 149
Qy 122 VMCTAFHCHNEETFLKVLVEYARHPVFYAFELLFFPAKVKYKAFTCCQAADKAACL 161
Db 150 VICEYQDNVSTFCHFIYSVARHPFLYAPAILSTFVDFEHALQSCCKSDVGACGLDTK 209
Qy 182 LDELDEGKASAKQRLKASLQKFGERAFAKAVARLSQRFPAKFAEYVSK-VTDLTKV 241
Db 210 EIVMREKAGSVYKQVFCGLIKQGDVRFQARQLIVLSQYKPAFSEVSKFVHDSIGV 269
Qy 242 HTECHGDLLJECADRDADLAKYICENQSSSSKKECEKEXPLLEKSHCIAVENDEMFAD 302
Db 270 HKECEGDMVECYDMARMMSNCSQCVFSGKIKDCCEKPIVERSQCINEAEFDEXPAD 329
Qy 302 LPSLAADVESEKQVCKNVAEAKQVFLGFLNVEYARRHPDYSVLLBLAKTYETTLKCC 362
Db 330 LPSLVEKYIECKVCKSPFAGHDAPFAEFVVEYRRHPFESQLIMRIAGYESLLEKCC 389
  
```

```

Qy 362 AAADPHCYAKVPBEPKPLVEEPQNLIKONCELFQEQSEYKFNALLVRYTKKYPQVSTP 421
Db 390 KTDNPAECYANAQOLNQH-KETQDVVKNTCDLLHDHGEADFLKSLILRY-KWNPQVPTD 449
Qy 422 TLVEVSRNLGVSKCKCKHPKAPKAPCAEDYLSVVLNQLCVLHEKTPVSDVTKCTTESL 481
Db 450 LLETKGKXMTTIGTKCCQLGDEDRMACSEGYLSVIHDTCKKQETTPINDKVSQCCSOLY 509
Qy 482 VNRRPCFSALEVDYTPKFEFNAETTFHADICTLSEKERCIKQTALVELVKKHKPKATK 541
Db 510 ANRRPCFTAMGVDTKYVPPPNPMFSEDEXLCSAPAEEREVGOMKLLINIKKQPMTE 569
Qy 542 EQLKAVMDOFAAFVEKCKCKADKCTCFAEBCCKVAASQAALGL 585
Db 570 EQIKTIADGFTAMVDKCKCKQSDINTCFEGGANLIVQSRATLGI 613

RESULT 14
FETA PANTR
ID FETA PANTR STANDARD; PRT: 609 AA.
AC Q28789;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-fetoprotein).
DE Pan troglodytes (Chimpanzee).
GS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_TaxID=9598;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=96032345; PubMed=7557431;
RA Nishio H., Gibbs P.E., Minghetti P.P., Zielinski R., Dugaiczak A.;
RT "The chimpanzee alpha-fetoprotein-encoding gene shows structural
RT similarity to that of gorilla but distinct differences from that of
RT human.";
RL Gene 162:213-220(1995).
CC 1- FUNCTION: BINDS COPPER, NICKEL, AND FATTY ACIDS AS WELL AS, AND
CC BILIRUBIN LESS WELL THAN, SERUM ALBUMIN. ONLY A SMALL PERCENTAGE
CC (LESS THAN 2%) OF THE HUMAN AFP SHOWS ESTROGEN-BINDING PROPERTIES.
CC 1- SUBUNIT: DIMERIC AND TRIMERIC FORMS HAVE BEEN FOUND IN ADDITION
CC TO THE MONOMERIC FORM.
CC 1- SUBCELLULAR LOCATION: Secreted.
CC 1- TISSUE SPECIFICITY: Plasma. Synthesized by the fetal liver and
CC yolk sac.
CC 1- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
CC 1- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

EMBL: U21916; AAA91641.1; -
 PIR: JC4258; JC4258.
 HSP: P02768; IE7B.
 DR InterPro: IPR000264; Serum_albumin.
 DR Pfam: PF00273; transport_prot; 3.
 DR PRINTS: PR00802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum_albumin; 1.
 DR SMART: SM00103; ALBUMIN; 3.
 DR PROSITE: PS00212; ALBUMIN; 2.
 KW Glycoprotein; Sulfation; Repeat; Metal-binding; Copper; Nickel;
 Signal.
 FT SIGNAL 1 18 BY SIMILARITY.
 FT CHAIN 19 609 ALPHA-FETOPROTEIN.
 FT DOMAIN 20 205 ALBUMIN 1.
 FT DOMAIN 212 397 ALBUMIN 2.

FT	DOMAIN	404	595	ALBUMIN 3.	
FT	METAL	22	22	COPPER AND NICKEL (BY SIMILARITY).	
FT	DISULFID	99	114	BY SIMILARITY.	
FT	DISULFID	123	124	BY SIMILARITY.	
FT	DISULFID	148	193	BY SIMILARITY.	
FT	DISULFID	192	201	BY SIMILARITY.	
FT	DISULFID	224	270	BY SIMILARITY.	
FT	DISULFID	269	277	BY SIMILARITY.	
FT	DISULFID	289	303	BY SIMILARITY.	
FT	DISULFID	302	313	BY SIMILARITY.	
FT	DISULFID	384	393	BY SIMILARITY.	
FT	DISULFID	416	462	BY SIMILARITY.	
FT	DISULFID	461	472	BY SIMILARITY.	
FT	DISULFID	485	501	BY SIMILARITY.	
FT	DISULFID	500	511	BY SIMILARITY.	
FT	DISULFID	538	583	BY SIMILARITY.	
FT	DISULFID	582	591	BY SIMILARITY.	
FT	CARBOHYD	42	42	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	251	251	N-LINKED (GLCNAC. . .) (POTENTIAL).	
SQ	SEQUENCE	609 AA;	68741 MW;	C012997CAD05672B CRC64;	
Query Match 40.4%; Score 1253.5; DB 1; Length 609;					
Best Local Similarity 40.1%; Pred. No. 2.4e-74;					
Matches 236; Conservative 115; Mismatches 231; Indels 7; Gaps 3;					
QY	3 HKSE----	VAHRPKDLGEENFKALVLIAPQYLCQGFEDHVKLVNTEVFAKTCVAD	57		
DB	22 HRNEVGYASILDSVQCTAEINLTDATFFAQFVQEAITYKEVSKVXDALTA-EKPTSD	81			
QY	56 SAENCKSLHTFGKLTQVATLETETGEMADCAKQEPERNECFLOHKDNP-NLPPLV	116			
DB	82 QSAGLEQLPAF-EELCREKEILEKYGH-SDCCSQSEEGRHNGFLAHKKPTASIPFFQ	140			
QY	117 RPENDVMCTAFHDEETLKKVLYEIAREHPYFVAPELLFPKRYKAFTCCCAADKAA	176			
DB	141 VPEPTVSCAEVEEETPMNPFVEIARRHFLVAPTILJWAARYDKIIPSCCKASHAVE	200			
QY	177 CLPKPLDELRDEGKASSAKQKLCASLQKGFGERAFKAWAVARLSQRPFKAEFAEVSRLVT	236			
DB	201 CFOTKAATVTKELRESSLLNQHACAVMKNFGTRTFQALTV-KLSQKTKVNFTEIQKLV	260			
QY	237 DLTKVHTCCGGDLLECDADRLAKYTCENQDISSKKKECKKPLEKSHCIAEVEND	296			
DB	261 DVARVHEHCRCRGDLVCLQDQGEKIMSY-CSQDD-LSNKITCECKLTTLERQCQIIHAEND	320			
QY	297 EMPADPLSLADQFVESKVCNRYABAKDVLGMEFYEVARRHPDYSVLLRLAKTYETT	356			
DB	321 EKPEGLSNLNRFGCRDNFNOFSQEKKIFLASPVHEYSRRHPQAVSVILRVAKGQEL	380			
QY	357 LEKCAAAADPHECYAKVDFEFPKPLVEEPQNLKQCELFQGLGEYKFNALLVRYTKVP	416			
DB	381 LEKCFQTEPLECQDKGEEELQKYIOESCALAKRSGLFQKLGVEYLLQNAFLVAYTKAP	440			
QY	417 QVSTTIVEVNRNLGVSKCKKHEAKMPCAEADYLSVLNQLCVLHEKTPNSCRVTKC	476			
DB	441 QJTSSELMAITKMAATAATCCQSEDKLLACGEGAADIIGHICIRHETTPVPGVQC	500			
QY	477 CTESLWNRPCFSALEVDETYVPKEFNAETTFTHADICTLSEKQKQKQATVVELVKHK	536			
DB	501 CTSSVNRPCFESSLVVETVVPFAFSDDKIFHKLQCAQVALQTMKQEFLLNLVQK	560			
QY	537 PKATKEQLKAVNDDFAAFVEKCKKADDDKTCFAEEGKKLVAAASQAALGL	585			
DB	561 PQITTEQLKAVIADFSGLLEKCKCQGEQEVCFABEGQKLSKTRAAALGV	609			

RESULT 15

FETA HUMAN

ID FETA HUMAN

AC F02471;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 29-FEB-2003 (Rel. 41, Last annotation update)

DE	Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-fetoprotein).
DE	APP.
GN	Homo sapiens (Human).
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1].
RN	SEQUENCE FROM N.A.
RX	MEDLINE=83273664; PubMed=6192439;
RA	Morinaga T., Sakai M., Wegmann T.G., Tamaoki T.;
RT	"Primary structures of human alpha-fetoprotein and its mRNA.";
RL	Proc. Natl. Acad. Sci. U.S.A. 80:4604-4608(1983).
RN	[2].
RN	SEQUENCE FROM N.A.
RX	MEDLINE=87185438; PubMed=2436661;
RA	Gibbs P.E.M., Zieliński R., Boyd C., Dugaiczak A.;
RT	"Structure, polymorphism, and novel repeated DNA elements revealed by a complete sequence of the human alpha-fetoprotein gene.";
RL	Biochemistry 26:1332-1343(1987).
RN	[3].
RN	SEQUENCE OF 1-28 FROM N.A.
RX	MEDLINE=93278385; PubMed=7684942;
RA	Mcvey C.H., Michaelides K., Hansen L.P., Ferguson-Smith X.,
RA	Tilghman S., Krumlauf R., Tuddenham E.G.D.;
RT	"A G->A substitution in an HNF I binding site in the human alpha-fetoprotein gene is associated with hereditary persistence of alpha-fetoprotein (HPAFP).";
RL	Hum. Mol. Genet. 2:379-379(1993).
RN	[4].
RN	SEQUENCE OF 429-556 FROM N.A.
RX	MEDLINE=83158778; PubMed=6187626;
RA	Beattie W.G., Dugaiczak A.;
RT	"Structure and evolution of human alpha-fetoprotein: deduced from partial sequence of cloned cDNA.";
RL	Gene 20:415-422(1982).
RN	[5].
RN	PARTIAL SEQUENCE OF 19-609.
RX	MEDLINE=91242409; PubMed=1709810;
RA	Pucci P., Siciliano R., Malorni A., Marino G., Tecce M.F.,
RA	Ceccarini C., Terrana B.;
RT	"Human alpha-fetoprotein primary structure: a mass spectrometric study.";
RL	Biochemistry 30:5061-5066(1991).
RN	[6].
RN	PRELIMINARY SEQUENCE OF 19-35.
RX	MEDLINE=77242506; PubMed=70228;
RA	Tactnin S., Hsu R., Heinrichson R.L., Miller J.B.;
RT	"Studies on human alpha-fetoprotein. Isolation and characterization of monomeric and polymeric forms and amino-terminal sequence analysis.";
RL	Biochim. Biophys. Acta 493:418-428(1977).
RN	[7].
RN	PRELIMINARY SEQUENCE OF 19-38.
RX	MEDLINE=78001760; PubMed=711198;
RA	Aoyagi Y., Ikenaka T., Ichida F.;
RT	"Comparative chemical structures of human alpha-fetoproteins from fetal serum and from ascites fluid of a patient with hepatoma.";
RL	Cancer Res. 37:3663-3667(1977).
RN	[8].
RN	PRELIMINARY SEQUENCE OF 19-39.
RX	MEDLINE=75018719; PubMed=4138095;
RA	Ruoslathi E., Pihko H., Vaheri A., Seppala M., Virolainen M.,
RA	Kontinen A.;
RT	"Alpha fetoprotein: structure and expression in man and inbred mouse strains under normal conditions and liver injury.";
RL	Johns Hopkins Med. J. Suppl. 3:249-255(1974).
RN	[9].
RN	GENE STRUCTURE.
RP	MEDLINE=85182629; PubMed=2580830;
RX	Sakai M., Morinaga T., Urano Y., Watanabe K., Wegmann T.G.,
RA	Tamaoki T.;
RT	"The human alpha-fetoprotein gene. Sequence organization and the 5'

RT flanking region.";
 RL J. Biol. Chem. 253:5055-5060(1985).
 RN [10].
 RP METAL-BINDING.
 RX MEDLINE=79001617; PubMed=80265;
 RA Aoyagi Y., Ikenaka T., Ichida F.;
 RT "Copper(II)-binding ability of human alpha-fetoprotein.";
 RL Cancer Res. 38:3483-3486(1978).
 RN [11].
 RP BILIRUBIN-BINDING.
 RX MEDLINE=80001710; PubMed=89600;
 RA Aoyagi Y., Ikenaka T., Ichida F.;
 RT "Alpha-fetoprotein as a carrier protein in plasma and its bilirubin-binding ability.";
 RL Cancer Res. 39:3571-3574(1979).
 RN [12].
 RP SULFATION.
 RX MEDLINE=86042625; PubMed=2414772;
 RA Liu M.C., Yu S., Sy J., Redman C.M., Lipmann F.;
 RT "Tyrosine sulfation of proteins from the human hepatoma cell line HepG2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).
 CC -!- FUNCTION: BINDS COPPER, NICKEL, AND FATTY ACIDS AS WELL AS, AND BILIRUBIN LESS WELL THAN, SERUM ALBUMIN. ONLY A SMALL PERCENTAGE (LESS THAN 2%) OF THE HUMAN AFP SHOWS ESTROGEN-BINDING PROPERTIES.
 CC -!- SUBUNIT: DIMERIC AND TRIMERIC FORMS HAVE BEEN FOUND IN ADDITION TO THE MONOMERIC FORM.
 CC -!- TISSUE SPECIFICITY: Secreted.
 CC -!- TISSUE SPECIFICITY: PLASMA. SYNTHESIZED BY THE FETAL LIVER AND YOLK SAC.
 CC -!- DEVELOPMENTAL STAGE: OCCURS IN THE PLASMA OF FETUSES MORE THAN 4 WEEKS OLD, REACHES THE HIGHEST LEVELS DURING THE 12TH-16TH WEEK OF GESTATION, AND DROPS TO TRACE AMOUNTS AFTER BIRTH. THE SERUM LEVEL IN ADULTS IS USUALLY LESS THAN 40 NG/ML. AFP OCCURS ALSO AT HIGH LEVELS IN THE PLASMA AND ASCITIC FLUID OF ADULTS WITH HEPATOMA.
 CC -!- PTM: INDEPENDENT STUDIES SUGGEST HETEROGENEITY OF THE AMINO-TERMINAL SEQUENCE OF THE MATURE PROTEIN AND OF THE CLEAVAGE SITE OF THE SIGNAL SEQUENCE.
 CC -!- PTM: SULFATED.
 CC -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M10949; AA451674.1; -
 DR EMBL; M10950; AA451675.1; -
 DR EMBL; V21514; CA424756.1; -
 DR EMBL; M16110; AA38754.1; -
 DR EMBL; Z19532; CA479592.1; -
 DR PIR; A26524; FPHJ.
 DR HSSP; PC2768; 1E7S.
 DR GlycoSuiteDB; P02771; -
 DR Siena-2DPAGE; P02771; -
 DR Genew; HGNC:317; AFP.
 DR MIM; 104150; -
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transport prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 2.
 DR Glycoprotein; Sulfation; Repeat; Metal-binding; Copper; Nickel;
 KW Signal; Polymorphism.
 FT SIGNAL; 18
 FT CHAIN; 19 609 ALPHA-FETOPROTEIN.
 FT DOMAIN; 20 205 ALBUMIN 1.
 FT DOMAIN; 212 397 ALBUMIN 2.

FT	DOMAIN	404	595	ALBUMIN 3. COPPER AND NICKEL.
FT	METAL	22	22	
FT	D-SULFID	99	114	
FT	D-SULFID	113	124	
FT	D-SULFID	148	193	
FT	D-SULFID	192	201	
FT	D-SULFID	224	270	
FT	D-SULFID	269	277	
FT	D-SULFID	289	303	
FT	D-SULFID	302	313	
FT	D-SULFID	384	393	
FT	D-SULFID	416	462	
FT	D-SULFID	461	472	
FT	D-SULFID	485	501	
FT	D-SULFID	500	511	
FT	D-SULFID	538	583	
FT	D-SULFID	582	591	
FT	CARBOHYD	251	251	
FT	VARIANT	570	570	
FT	SEQUENCE	609 AA;	68677 MW;	
SC	SEQUENCE	609 AA;	68677 MW;	
Query Match		40.3%;	Score 1249.5;	DB 1; Length 609;
Best Local Similarity		39.9%;	Pred. No. 4.4e-74;	
Matches		235;	Conservative 116;	Mismatches 231; Indels 7; Gaps 3;
QY	3 HKSE----	VAHRFKDLGEENFKALVLI	FAQYLOQCPFDHVKLVN	VEFEAKTCVADE 57
DB	22 HRNEYGASILDSYQCTA	ISLADLATIFFAQVQVQ	QATYKEVSKMVKD	ALTAIEKPTGE 81
QY	58 SARNDKSLHTJFGDKLCT	VATLRETYGEMADCCAKQ	EPERNECFLOHXDNP	NILPRLV 116
DB	82 QSSGLENQLPAFLBELCH	EKEILEKVGH-SDCCSQS	SEEGRNCFLAHKKPT	PASIPLEQ 140
QY	117 RPEVDVMTAFHONBET	FKKLYIEIARRHPYAP	LELLFFAKRYKAA	TECCQAADKAA 176
DB	141 VPEPVTSCAEYEDRET	FMNKFIVEIARRHPFL	YAPTILLMAARYDKI	IPSCCKAENAVE 200
QY	177 CLLPKLDELDEGKASSA	KQRLCASLQKFGERAKA	VARUSQRPFAEAEVSK	LVT 236
DB	201 CFQTKAATVTKELRES	SLNQHACAVMKVGTFT	FOAITVKLSQKFTKVN	FTFIOKLVL 260
QY	237 DLTKVHTECHGDLLECA	DDRADLAKYICENQDSI	SSKKECEKPLEKSHC	IAEVEND 296
DB	261 DVAHVHEHCCRGDVLDC	LQDGEKIMSYCSQD	TLNKNITECKLTJER	GCIIHAEND 320
QY	297 EMPADPLSLAADFESK	VDYCKNYAEAKDVLGM	FLMFLYEAAREHPD	YSVVLRLAKTYET 356
DB	321 EKPEGLSPNLRFLGDR	DFNQFSGGKNITFLA	SVHEYSRRHPOLAV	SVTLRVAKGYOEL 380
QY	357 LEKCCAAAPHECYAKV	DFEFKPLVEEPQNLK	IKONFELFQLGEYK	FQNALJLVRYTKVP 416
DB	381 LEKCFOTENPLECODK	GEEELQKYIGESQAL	AKSSCGLFKQKGEY	YLQNAFJVAITYKAP 440
QY	417 QVSTPTLVEVSRNLG	VGSKCKKHPAKRYPC	AEADLSVVLNQLCVL	HEKTPVSRVTKC 476
DB	441 QLTSSSELVAITRKMA	ATAATCCQLSEDKL	ACGGAADIIGHLCI	RRHEMTFVNPVGQC 500
QY	477 CTESLVNRRPFSAL	EVDETYVFEKENAE	FTFHADICTLSEKER	QIKKOTALVELVKHK 536
DB	501 CTSSYANRRPFCSS	LVVETVVPFPAFSD	DKFIFHKOLCOAQ	GQVVALOTMKEFLINLVKQK 560
QY	537 PKATKEOLKAVMD	DFAAFEVKCKCKAD	KETCFAEEGKLV	AAASQAALGL 585
DB	561 PQTTEOLEAVIAD	FSGLLEKCKCQ	QEQEVCFAE	EGQKLSIKTRALGV 609

Search completed: October 27, 2003, 15:19:12
 Job time : 27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein: - protein search, using sw model

Run on: October 27, 2003, 15:15:23 ; Search time 106 Seconds
(without alignments)
1424.156 Million cell updates/sec

Title: US-09-833-117-18

Perfect score: 3103

Sequence: 1 DAHSEVAHRFKDLGSENFK.....TCFAEKGKLVAAQAALG: 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SFTRMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2504	80.7	609	5 Q95VB7	Q95VB7 schistosoma
2	2374	76.5	609	11 C8CTH3	C8CTH3 mus musculus
3	2336	75.3	576	1: Q8CTC7	Q8CTC7 mus musculus
4	1865.5	60.1	396	4 Q8IUK7	Q8IUK7 homo sapien
5	1295.5	41.7	527	13 Q8JIA9	Q8JIA9 sphenodon p
6	1242	40.0	609	6 Q8MJU5	Q8MJU5 canis famil
7	1242	40.0	626	13 Q8JIA9	Q8JIA9 ambystoma m
8	1218.5	39.3	610	6 Q8MJ76	Q8MJ76 sus scrofa
9	1087	35.0	624	13 C8UW06	C8UW06 ambystoma t
10	1082	34.9	605	11 C8BK65	C8BK65 mus musculus
11	1076	34.7	605	11 C8BK56	C8BK56 mus musculus
12	1045	33.7	400	13 Q8JIA7	Q8JIA7 sphenodon p
13	955	30.8	603	13 Q9YGH6	Q9YGH6 rana shqipe
14	928.5	29.9	614	13 Q9L134	Q9L134 naja naja
15	888	28.6	406	13 Q8JIA8	Q8JIA8 hepiodactyl
16	739	23.6	235	11 Q8CG74	Q8CG74 mus musculus

17	713	23.0	47.7	11 Q8R0J9	Q8R0J9 mus musculi
18	376.5	12.1	484	13 Q9W6F5	Q9W6F5 gallus gall
19	373	12.0	476	11 Q9CY31	Q9CY31 mus musculu
20	372	12.0	476	11 Q9IXG1	Q9IXG1 mus musculu
21	365	11.8	551	13 Q42279	Q42279 petromysc
22	331	10.7	122	13 Q90WZ8	Q90WZ8 iarus argen
23	290	9.3	123	13 Q90WZ6	Q90WZ6 poephila gu
24	264	8.5	135	11 Q63205	Q63205 rattus ncrv
25	186	6.1	1723	2 Q9JMX8	Q9JMX8 heliobacte
26	184	5.9	1819	16 Q9ZLV0	Q9ZLV0 heliobacte
27	184	5.9	1927	16 Q25262	Q25262 heliobacte
28	162.5	5.2	44	6 Q95MC2	Q95MC2 equus cabal
29	162.5	5.2	680	5 Q9V6S8	Q9V6S8 drosophila
30	161	5.2	1079	3 Q96V11	Q96V11 pneumocysti
31	156	5.0	1026	3 Q74669	Q74669 pneumocysti
32	154.5	5.0	3843	5 Q9USD0	Q9USD0 drosophila
33	153.5	4.9	661	5 Q8MS79	Q8MS79 drosophila
34	153.5	4.9	3843	5 Q9VJ94	Q9VJ94 drosophila
35	153	4.9	62	6 Q8MIL1	Q8MIL1 sus scrofa
36	150	4.8	1065	3 Q01828	Q01828 pneumocysti
37	149	4.8	40	6 Q9TRA5	Q9TRA5 oryctolagus
38	148	4.8	1028	3 Q74668	Q74668 pneumocysti
39	144.5	4.7	8749	4 Q8NF91	Q8NF91 homo sapien
40	142.5	4.6	2756	10 Q9LJ60	Q9LJ60 arabidopsis
41	141.5	4.6	1560	5 Q26644	Q26644 strongyloce
42	138.5	4.5	1069	3 Q96V12	Q96V12 pneumocysti
43	137.5	4.4	1348	16 Q8YK55	Q8YK55 anabaena sp
44	135.5	4.4	2841	5 Q8MLJ9	Q8MLJ9 drosophila
45	135.5	4.4	2931	5 Q9W2C6	Q9W2C6 drosophila

ALIGNMENTS

RESULT 1

Q95VB7 PRELIMINARY; PR: 608 AA.
AC Q95VB7; 19, Created;
ET 01-DEC-2001 (TRENBLrel. 19, Last sequence update);
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update);
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update);
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6483;
RN [1]_TaxID=6483;
RP SEQUENCE FROM N.A.
RA Osmani A., Asahi H., Stadelker M.J., LoVerde P.T.;
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg
component in murine infection with Schistosoma mansoni";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF418550; AAL08579.1; ...
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PDC02486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 68225 MW; E5EABB28E1C66E54 CRC64;

Query Match 80.7%; Score 2504; DB 5; Length 608;
Best Local Similarity 76.3%; Pred. No. 1.7e-188;
Matches 445; Conservative 79; Mismatches 59; Indels 0; Gaps 0;

QY	1	DAHSEVAHRFKDLGSENFKALVIAFAVLQCPFFDHVKLVNEVTEFAKTCVADESAAE 60
DB	25	DAHSEIAHREKDLGEQHPKGLVLIASFQLOKQPYEHRVKLVNEVTFDKTCVADESAAE 84
QY	61	NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPLRLVPRPEV 120
DB	85	NCCKSLHTLFGDKLCAIPTLRLDSYGLADCCAKKEPRNECFLOHKDDHPLPFPVRPDA 144

```
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCOQAAQAACTLP 180
DB 145 EAMCTSFQENAVFMGHYLVHEVARRHPYFYAPPELLFFAYAEKYSAINTECCGEADKAACITP 254
QY 181 KJDELRDEGKASSAKORUKKASLOKGERAFKAWAVARLSQRPKAEFAEVSQVTDLT 240
DB 205 KJDAKKEKALASSVNGRLKSSQRFQORAFKAWAVARMSQRPKADFAEITKATDCLK 264
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 265 LTBECCHGDLLECADRAELAKYICENQASISSKLOACCKEVLKXSHCLSEVENDELPA 324
QY 301 DLPSLAADFVESKDYCKNYAEAKOVFLGMEFYAEARRHPDYSVVLLRLAKTYETLEKX 360
DB 325 DLPSLAADFVECKEVCNKYAEAKOVFLGTLFYEARRHPDYSVALLRLAKKYEATLEK 384
QY 361 CAADPHCEYAKVDFEFPKPLVEBPONLJKNCLEFEOLGEVKFCNALVRYTKVPQVST 420
DB 385 CAEADPSACYGKVJDEFQPLVEBPKNLVKANCELFELKLGEGYFQNALVRYTKAPQVST 444
QY 421 PTLVESVRNLGKVGSKCKEPEAKMPCAEADYLSWJNQLCVLHEKTPVSDRVTKCTES 480
DB 445 PTLVESVRNLGKVGSKCVLPPEAGLSQVEDYISAILNRVCVLHEKTPVSEQVTKCTGS 504
QY 481 LVNRRPCFSALEVDITYVPKEFNASTFTPHADICTLSEKERQIKKQTAJLVKHKPKAT 540
DB 505 VERRPCFSALPVZETVYVPKEFAETTFHSDICTLPEKEQIKKQTAJLVKHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKKADDEKTCFAEBGKKLVAAASQAAL 583
DB 565 GPQJRTVLGEFTAFEDKCKRAEDKAEFSEDDGPKLVASSQAAL 607

RESULT 2
Q8C7H3 PRELIMINARY; PRT; 608 AA.
ID Q8C7H3
AC Q8C7H3:
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Albumin 1 (Fragment);
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
DR Nature 420:563-573(2002).
DR EMBL; AK050248; BAC34145.1; -.
FT NON_TER
SQ SEQUENCE 608 AA; 68722 MW; 292F603BED3A61B4 CRC64;

Query Match 76.5%; Score 2374; DB 11; Length 608;
Best Local Similarity 72.2%; Pred. No. 2.7e-178;
Matches 421; Conservative 80; Mismatches 82; Indels 0; Gaps 0;

QY 1 DAHSEVARRFKDGEENFKALVLIAPAQYLOQCPPEFHVKLVNEVTEFAKTCVADSESAE 60
DB 25 EAHKSEIARRNDYGEQHFKEGLVLIAPFQYLOKCSYDEHAKLVQEVTDFAKTCVADSESA 84
QY 61 NCKDSLHTLFGDKLCTVATLSETYGMADCCAKQEPERNECFLOHKQDNPNLRLVPEV 120
DB 85 NCKDSLHTLFGDKLCAIENLRENYGELADCTCKQEPERNECFLOHKQDNPNLPPFERPEA 144
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCOQAAQAACTLP 180
DB 145 EAMCTSFQENAVFMGHYLVHEVARRHPYFYAPPELLFFAYAEKYSAINTECCGEADKAACITP 204
```

```
QY 181 KJDELRDEGKASSAKORUKKASLOKGERAFKAWAVARLSQRPKAEFAEVSQVTDLT 240
DB 205 KJDAKKEKALASSVNGRLKSSQRFQORAFKAWAVARMSQRPKADFAEITKATDCLK 264
QY 241 VHTTECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 265 LTBECCHGDLLECADRAELAKYICENQATSSKLOACCKEVLKXSHCLSEVENDELPA 324
QY 301 DLPSLAADFVESKDYCKNYAEAKOVFLGMEFYAEARRHPDYSVVLLRLAKTYETLEKX 360
DB 325 DLPSLAADFVECKEVCNKYAEAKOVFLGTLFYEARRHPDYSVALLRLAKKYEATLEK 384
QY 361 CAADPHCEYAKVDFEFPKPLVEBPONLJKNCLEFEOLGEVKFCNALVRYTKVPQVST 420
DB 385 CAEANPPACYGKVLAEFQPLVEEPKNLVKNCGLYKLGEGYFQNALVRYTKAPQVST 444
QY 421 PTLVESVRNLGKVGSKCKEPEAKMPCAEADYLSWJNQLCVLHEKTPVSDRVTKCTES 480
DB 445 PTLVESVRNLGKVGSKCVLPPEAGLSQVEDYISAILNRVCVLHEKTPVSEQVTKCTGS 504
QY 481 LVNRRPCFSALEVDITYVPKEFNASTFTPHADICTLSEKERQIKKQTAJLVKHKPKAT 540
DB 505 VERRPCFSALPVZETVYVPKEFAETTFHSDICTLPEKEQIKKQTAJLVKHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKKADDEKTCFAEBGKKLVAAASQAAL 583
DB 565 ABQJRTVLGEFTAFEDKCKRAEDKAEFSEDDGPKLVTRCKDTL 607

RESULT 3
Q8C7C7 PRELIMINARY; PRT; 576 AA.
ID Q8C7C7
AC Q8C7C7:
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Albumin 1 (Fragment);
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
DR Nature 420:563-573(2002).
DR EMBL; AK050644; BAC34360.1; -.
FT NON_TER
SQ SEQUENCE 576 AA; 65002 MW; F85733E99AE37F04 CRC64;

Query Match 75.3%; Score 2336; DB 11; Length 576;
Best Local Similarity 72.2%; Pred. No. 2.5e-175;
Matches 415; Conservative 79; Mismatches 81; Indels 0; Gaps 0;

QY 9 HRFKDLGEENFKALVLIAPAQYLOQCPPEFHVKLVNEVTEFAKTCVADSESAENCKSLHT 68
DB 1 NRYVDLGEQHFKEGLVLIAPFQYLOKCSYDEHAKLVQEVTDFAKTCVADSESAENCKSLHT 60
QY 69 LFGDKLCTVATLSETYGMADCCAKQEPERNECFLOHKQDNPNLRLVPEVVMCTAFH 128
DB 61 LFGDKLCAIENLRENYGELADCTCKQEPERNECFLOHKQDNPNLPPFERPEAEAMCTSEK 120
QY 129 DREETFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCOQAAQAACTLP 188
DB 121 ENPTFMGHYLVHEVARRHPYFYAPPELLFFAYAEQYNEILTQCCAEADKESCLT 180
QY 189 GXASSAKORUKKASLOKGERAFKAWAVARLSQRPKAEFAEVSQVTDLT 248
DB 181 ALVSSVRQMKSSQKQGERAFKAWAVARLSQTFNADFAEITKATDCLK 240
```



```

QY 249 DLLECAADDRADJAKY:CEQODSISSKLECCERPLLEKSHCIAEVENDEMPAELSLAAD 308
DB 241 DLLECAADDRADJAKY:CEQODSISSKLECCERPLLEKSHCIAEVENDEMPAELSLAAD 300
QY 309 FVSEKDVCKNYAEAKDVFGLMFLEYEARHPDYVSUULLRLAKTYETTLEKCAAAAPHE 368
DB 302 FVEDQEVCKNYAEAKDVFGLMFLEYEARHPDYVSUULLRLAKTYEATLEKCAEAPPA 360
QY 369 CYAKVDFEPLVSEPNLIKONCELFQELGEYKFNALLVRYTKKVPQVSTPTLVEVSR 428
DB 361 CYGTVLAEFQPLVEEPKLVK:NCCLYELKGLBYGFQNAIVRYTKAPQVSTPTLVEAAR 420
QY 429 NLGKVGSKCKHPEAKRMPCAEADYLSVNLQCLVLEKTPVSDRVTKCCCTESLVNRRPCF 488
DB 421 NLGRVGTCKCTLPEDGRPCVDEYLSAUNRVCLLEKTPVSHVTKCCSGSUVERRPCF 480
QY 489 SALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVHKHPKATKEOLKAYM 548
DB 481 SAUTVDETYVPKEFNAETFTFHSDICTLPEKEKQIKKQATALVELVHKHPKATAEOLKAYM 540
QY 549 DFEAAAFVEKCKADDDKTCFAEEGKLVAAASQAAL 583
DB 541 DFEAQFLDTCCKAAGKOTCFSTEGPNLVTRCKDAL 575

```

RESULT 4

```

Q81UK7 PRELIMINARY; PRT: 396 AA.
AC Q81UK7
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Similar to serum albumin precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RJ Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035969.1;
SQ SEQUENCE 396 AA; 45153 MW; 756519C096463A9B CRC64;

```

```

Query Match 60.8%; Score 1865.5; DB 4; Length 396;
Best Local Similarity 63.8%; Pred. No. 1.6e-138;
Matches 372; Conservative 0; Mismatches 0; Indels 213; Gaps 1;

```

```

QY 1 DAHSEVAHRFKDLGRENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESA 60
DB 25 DAHSEVAHRFKDLGRENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESA 84
QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVREPV 120
DB 85 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVREPV 144
QY 121 DVMTAFHDNEETFLKXYLVEIARRHPDYFAPPELLFFAKRYKAAFTCCCAADKACCLP 180
DB 145 DVMTAFHDNEETFLKXYLVEIARRHPDYFAPPELLFFAKRYKAAFTCCCAADKACCLP 163
QY 181 KLDELDRDEGKASSAKORLKASLQKFGERAFKAWAVARLSQRPFAEFAVSKLVDTLTK 240
DB 164 ----- 163
QY 241 VHTCCCHGDLLECAADDRADJAKYICENQDSISSKLECCERPLLEKSHCIAEVENDEMPA 300
DB 164 ----- 163
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYEARHPDYVSUULLRLAKTYETTLEK 360
DB 164 ----- 163
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYEARHPDYVSUULLRLAKTYETTLEK 360
DB 164 ----- 163

```

```

QY 361 CAADPHCEYAKVDEPKPLVSEPNLIKONCELFQELGEYKFNALLVRYTKKVPQVST 420
DB 172 CAADPHCEYAKVDEPKPLVSEPNLIKONCELFQELGEYKFNALLVRYTKKVPQVST 231
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEADYLSVNLQCLVLEKTPVSDRVTKCCCTES 480
DB 232 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEADYLSVNLQCLVLEKTPVSDRVTKCCCTES 291
QY 481 LVNRRPFCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVHKHPKAT 540
DB 292 LVNRRPFCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVHKHPKAT 351
QY 541 KEOLKAYMNDFFAAAFVEKCKADDDKTCFAEEGKLVAAASQAALGL 585
DB 352 KEOLKAYMNDFFAAAFVEKCKADDDKTCFAEEGKLVAAASQAALGL 396

```

RESULT 5

```

Q8JIA9 PRELIMINARY; PRT: 527 AA.
AC Q8JIA9
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Serum albumin (Fragment).
OS Sphenodon punctatus (Hastaria) (Tuatara).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
OX NCBI_TaxID=8508;
RN [1]
RP SEQUENCE FROM N.A.
RA Vercalci V.J., Brennan S.O., George P.M., Chambers G.K.;
RT "Partial coding sequence for Sphenodon punctatus 68 kDa albumin.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375971; AAM46104.1;
DR IPI:IPRO; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot; 3.
DR PRINTS; PR03802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
FT NON TER 1
SQ SEQUENCE 527 AA; 59711 MW; C62B799E387F5929 CRC64;

```

```

Query Match 41.7%; Score 1295.5; DB 13; Length 527;
Best Local Similarity 45.9%; Pred. No. 1.4e-93;
Matches 240; Conservative 92; Mismatches 188; Indels 3; Gaps 3;

```

```

QY 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVREPV 120
DB 5 CJKSLDTIFLDEICHEEGFAAKY-DLAACCAKAEVERKECLLAHKNATPGFAPQRPFI 63
QY 121 DVMTAFHDNEETFLKXYLVEIARRHPDYFAPPELLFFAKRYKAAFTCCCAADKACCLP 180
DB 64 EVSKCLYQDDRLTLGLNVIYEVARRHPYLVPPVFATASLYDEALKTCCTADKATCFHP 123
QY 181 KLDELDRDEGKASSAKORLKASLQKFGERAFKAWAVARLSQRPFAEFAVSKLVDTLTK 240
DB 124 RIPLIEIYLVKMSNGIOENTCGILKKFGERTLAKTLQVMSOKPFPADATINKLVEDITH 163
QY 241 VHTCCCHGDLLECAADDRADJAKYICENQDSISSKLECCERPLLEKSHCIAEVENDEMPA 300
DB 184 MHTCCCHGDLLECAADDRADJAKYICENQDSISSKLECCERPLLEKSHCIAEVENDEMPA 243
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYEARHPDYVSUULLRLAKTYETTLEK 360
DB 244 DLSEIAEYEDPHVCDHLAKEQDAFLAKFLYYSRRHPDLSTQILLGVGKGYQELLERC 303
QY 361 CAADPHCEYAKVDEPKPLVSEPNLIKONCELFQELGEYKFNALLVRYTKKVPQVST 420
DB 304 CKTDNPECYQGAADLKGHIAQFQELVQNCDDYNTLGGYLFHALLIRYTKRMPGLTS 363

```

```

QY 421 PTLVEVRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHETPTVSDRVTCKCTESLVNRRPCFSALEVD 480
D 421 PTLVEVRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHETPTVSDRVTCKCTESLVNRRPCFSALEVD 480
D 364 BELIFYTR-ITKAASRCCEVSVDKLPCTEGYVDFVLGQICQHRSS-INVWQCQCSNS 422
QY 481 LVNRRPCFSALEVDYTVPKPENAETFTPHADICTLSEKERQIKQTALVELVKKHPKAT 540
D 481 LVNRRPCFSALEVDYTVPKPENAETFTPHADICTLSEKERQIKQTALVELVKKHPKAT 540
D 423 YALRSLCITSLGGDEKPEIEFSADLTFPHEDJCHAAQDKLQBRKQOMIVNLVYKHPN- 482
QY 541 KEQLKAVYCCFAAFVEKCCKADKCTCFABEGKLVAAASQAL 583
D 541 KEQLKAVYCCFAAFVEKCCKADKCTCFABEGKLVAAASQAL 583
D 483 KEQLQTVFGGFTKMTKCCCKAEDEHAGCGEGPKVAESQAL 525

```

RESULT 6

```

Q8MJU5 Q8MJU5 PRELIMINARY; PRT; 609 AA.
AC Q8MJU5:
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Alpha-fetoprotein.
GN AFP
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
CX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Furutachi M., Neo S., Hisasue X., Tsuchiya R., Watanabe M.,
RA Hashizaki K., Hisamatsu S., Kamada T.;
RI "Canine alpha-fetoprotein cDNA."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB089789; BAC07513.1; -
DR InterPro: IPR000264; Serum albumin.
DR Pfam: PF00273; transport_2foc; 3.
DR PRINTS; PRO0802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 609 AA; 68782 MW; BE4B8250C5AF2AFC CRC64;

```

```

Query Match 40.0%; Score 1242; DB 6; Length 609;
Best Local Similarity 40.5%; Pred. No. 2.7e-89;
Matches 231; Conservative 133; Mismatches 225; Indels 2; Gaps 2;

```

```

QY 16 BENFKALVLIAPCYLOQCPEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC 75
D 40 EMNLVCLATIFFAGFQVQETKYKSKVKXDIITVIEKSTGSEQPGGLEQLPAPLEIIC 99
QY 76 TVATLRETYGEMADCCAKQEPERNECFLOHKD-DNPNLPLRVPEVDVMTCTAFHKEET 134
D 100 HEKISEIKYG-LADCCSQREBERENCFLAHKKAAPPSPFQVAEPVTSCKAYEENRDMF 159
QY 135 LKKYLYETARRHPYFAPELLFFAKRYKMAFTPECQAAQCAACLLPKDLERDEGRASSA 194
D 159 MNRVYIETARRHPYFAPELLFFAKRYKMAFTPECQAAQCAACLLPKDLERDEGRASSA 194
QY 195 KQRKCKASLQKFGERAFKAWAVARLSQRFPAKAEVSKLVTLTKVHTTECHDGLLECA 254
D 219 LNQHICAMRNFGPRTFRAITVTKLSQKFSKANTETQKLVDAVHIECCRGVLECL 278
QY 255 DDRADLAKYICENQCSISSKLKECEKPLEKSHCIAEVENDEMPADLPALADFVSKD 314
D 279 ODGEKIMSYICQDILSSKIDACKLPLELGOCIIHAENDGKPEGLSPNLNRFLEED 338
QY 315 VCKYVAEKQVFLGMFLYEYARRHPDYSVLLSLAKTYETITTEKCCAAADPHECYAKVF 374
D 339 FNQFSREKOLFMARFTVEYSRRHTKLAVPVVLVAKGYQELLESKCSQSENPECCQKGE 398
QY 375 DEFKPLVEPQNLKONCELPQELQGEKFNALLVRVTKVPQVSTPTFVSVSNLKGK 434
D 399 EELEKYIGESQALAKRSCGLFQKLGVEYLVQNAFLVATYTKAPQLTPPELKAFTKMTAA 458

```

```

QY 435 SKCKCHPEAKRMPCAEDYLSVVLNQLCVLHETPTVSDRVTCKCTESLVNRRPCFSALEVD 494
D 435 SKCKCHPEAKRMPCAEDYLSVVLNQLCVLHETPTVSDRVTCKCTESLVNRRPCFSALEVD 494
D 459 ATCCCLSDRCLACGEGAADLIIGQLCIRHETPTNPGVGCCSSSYANRRPCFSSLVVD 518
QY 495 ETVVPKPENAETFTPHADICTLSEKERQIKQTALVELVKKHPKATKEQLKAVYCCFAAF 554
D 495 ETVVPKPENAETFTPHADICTLSEKERQIKQTALVELVKKHPKATKEQLKAVYCCFAAF 554
D 519 ETVVPKPENAETFTPHADICTLSEKERQIKQTALVELVKKHPKATKEQLKAVYCCFAAF 578
QY 535 VEKCCKADKCTCFABEGKLVAAASQAL 585
D 535 VEKCCKADKCTCFABEGKLVAAASQAL 585
D 579 LEKCCGQEQEACFEFGPKLISKTRALGV 609

```

RESULT 7

```

Q8JW05 Q8JW05 PRELIMINARY; PRT; 626 AA.
AC Q8JW05:
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Serum albumin precursor.
GN A1B
OS Ambystoma maculatum (spotted salamander)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandridae; Ambystomatidae;
OC Ambystoma.
CX NCBI_TaxID=43114;
RN [1]
RP SEQUENCE FROM N.A.
RA TSSJUELiver;
RA Haverfield E.V., Uzzell T., Spolsky C.M., Bazartseren B.;
RI "Serum albumin of the mole salamanders Ambystoma maculatum and
RT Ambystoma texanum."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217183; AAL56646.1; -
DR InterPro: IPR000264; Serum albumin.
DR Pfam: PF00273; transport_prot; 3.
DR PRINTS; PRO0802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
KW Signal.
FT CHAIN 25 626 SERUM ALBUMIN.
SQ SEQUENCE 626 AA; 70677 MW; 9D66F57F174AC23F CRC64;

```

```

Query Match 40.0%; Score 1242; DB 13; Length 626;
Best Local Similarity 40.1%; Pred. No. 2.8e-89;
Matches 237; Conservative 110; Mismatches 238; Indels 6; Gaps 3;

```

```

QY 1 DAHSEVAHRFKDL----GEENFKALVLIAPCYLOQCPEDHVKLVNEVTEFAKTCTVAD 56
D 26 EGVNDNFPHLIGDLIPMGVDNSKGLVLAASOMLPFCPEEHLQVDEVDVMQIADLCAK 87
QY 57 ESAENCDKSLHTFGKCLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQD-PNLPL 115
D 88 ASHANCASPMITLIJDELCKPENAEKYPFHQECCKDEPERSHKCFVEHKNMHPHETKY 147
QY 116 VRPEVDVMTCTAFHNEETFKKLYEYIARRHPYFAPELLFFAKRYKMAFTPECQAAQCA 175
D 148 VRPAPEQICQKHAENRGPPLARYIIMLAIGHPHMVIIPALGPAQRPDGLVSHCCQDVET 207
QY 176 A-CLLPKLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRFPAKAEVSKL 234
D 208 GCFNDKMPHQEQVEYVVCALQKHCNYILQDFKRALTAAYKAVQASQKFFLASFENVQ 267
QY 235 VTDLTKVHTTECHDGLLESCADRADLAKYICENQCSISSKLKECEKPLLEKSHCIAEVE 294
D 268 VPDTVHJHOTCCGGGMACMLERMKLTAKICEKQDELATHLKECCDKPILLERSACIIRP 327
QY 295 NDEMPADLPALADFVSKDVCNKVAEAKDVLGMFLYEYARRHPDYSVLLSLAKTYE 354
D 328 NQOKPADLSPKVPHYIIDDPEVCCKLYTEGGDTFMGRFLYECARRHODYSPENLJRMGSGY 387

```

```

QY 355 TTLEKCCAAADPHECYAKVFDKPLVEEPONLIKONCELFEBOLGEYKFKONALLHYPTKK 414
Db 388 EFLKCCAAAGHNECAKTESKKEIESSVTJKTNCAGALDKKSLYFONLLIFKIVAR 447
QY 415 VPOVSTPTLVEVRNRLKGVSKCKPKEAKRMPCAEDY:SWJNQLCVLHEKTPVSDRYT 474
Db 448 MPALSEQSLARITKSMITIGEKCHRPEDQOMTCSEGLGIVGOICMKCKTTPUNEKVA 507
QY 475 KCTESLUNRRPFCFSALEVEDTVVPKEFNAETTFHADICTLSEKRO:KKQTALVELJYK 534
Db 508 QCCSHLSSTPFCFSALEVEDTVVPPLSVASFNFDLCTTSEPPQSKQVFLIRLMK 567
QY 535 HKPKATKEQLKAVMDPFAEVKCCXADDKETCFABEGKKVAASOALGL 585
Db 568 QYFMTDEQLKTCVNVFVPMVQCCXADNHNCFALGAKLIDACKAILAV 618

RESULT 8
Q8MJ76 PRELIMINARY; PRT; 610 AA.
AC Q8MJ76
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE Alpha-fetoprotein.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim J.G., Nonneman D., Vallet J.L., Christenson R.K.:
RT "Mapping of the porcine alpha-fetoprotein (AFP) gene to SSC8.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF517776; AAM66110.1; -.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 610 AA; 68624 MW; C985BEAD44963D5E CRC64;

Query Match: 39.3%; Score 1218.5; DB 6; Length 610;
Best Local Similarity 39.7%; Pred. No. 1:9e-87;
Matches 227; Conservative 117; Mismatches 225; Indels 3; Gaps 3;

QY 16 EENFKALVJIAFAQYLQCCPFEDHVKLVNVEVTEFAKTCVADESANCKSLHTLFCKJC 75
Db 40 EMLVLQATIFFAQVQGEATYKEVNCVMKDLTVIEKSTGSECPAGCLENGVSVFLEETC 99
QY 76 TVATLRETYGEMADCCAKQEPBERNECFLOHKKDNP-NLPLRVPEVDVMTAFPHDNEE 134
Db 100 HEEBIEPKYG-LSHCCSQSGSEERHNCFLARKKAAPASIPFPFPEVPTSCKAYEENRETF 158
QY 135 LKXYLYEIAARRHPYFAPELLTFKRYKAAFTCCQAADKAAACLLPKLDELHDEGRKASA 194
Db 159 MTRVYIEIARRHFLPYATILSLAAQYDKIIPCCKAENAVECFQTKAASIKELRESSL 218
QY 195 KQRLKASLQKGERAFKAVARLQRFPAKFAEVSKLVDTLTKVHTTCCCHGDLLECA 254
Db 219 LNOHMTVMRQFGARTTRAITVTLSQKFPKANFTETQKLVLDVAHIHEECRGNVLECL 278
QY 255 DORADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDENPADLPSLAADPVEK 313
Db 279 QDAERVVSVYCSQODTLSSKIAESCCKLPTTLELGGCIIHAENDKXPEGLSPNLNRELGR 338
QY 314 DVCKGVAEADVFLGMLFELYEARRRHPDYSVLLLRALAKTYETTTLEKCCAAADPHECYAKV 373
Db 339 DFNQLSREKDLNARTTYSRRHPKLVPTVILRVAKGYQELLEKCSOSRPLECQDKG 398
QY 374 DDEFKPLVEEPONLIKONCELFEBOLGEYKFKONALLHYPTTKVQVSTPTLVEVSRNLGKV 433

```

```

Db 399 BEELEKUIOESQALAKSCGFLQKLGYYLQNAFLVAYTKRAPOLTPPELVALTRKMAT 458
QY 434 GSKCKPKEAKRMPCAEDY:SWJNQLCVLHEKTPVSDRYTCKCTESLVNRRPFCFSALEV 493
Db 459 GAACCHLSEDRQLACGEGADLIIGQLCIRHEEMPINPGVGGCCTSSYANRRPCFSSLV 518
QY 494 DETVVPKEFNAETTFHADICTLSEKRO:KKQTALVELJYKHPKATKEQLKAVMDPFA 553
Db 519 DETVVPKEFNAETTFHADICTLSEKRO:KKQTALVELJYKHPKATKEQLKAVMDPFA 578
QY 554 FVEKCCXADDKETCFABEGKKVAASOALGL 585
Db 579 LLEKCCQCGQGEVCFABEGFALISKTRASLOV 610

RESULT 9
Q8UM06 PRELIMINARY; PRT; 624 AA.
AC Q8UM06
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE Serum albumin precursor.
OS Ambystoma texanum (Smallmouth salamander).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=8304;
RN [1]
RP SEQUENCE FROM N.A.
RA Haverfield E.V., Urzell T., Spolsky C.M., Bazartseren B.;
RT "Serum albumin of the mole salamanders Ambystoma maculatum and
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217182; AAL56645.1; -.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 1.
KW Signal.
FT SIGNAL. 1 24 POTENTIAL.
FT CHAIN. 25 624 SERUM_ALBUMIN.
SQ SEQUENCE 624 AA; 70321 MW; DE08533BF4953EF7 CRC64;

Query Match: 35.0%; Score 1087; DB 13; Length 624;
Best Local Similarity 37.7%; Pred. No. 4:3e-77;
Matches 214; Conservative 102; Mismatches 240; Indels 12; Gaps 5;

QY 14 LGBENFALVUIAFAQYLQCCPFEDHVKLVNVEVTEFAKTCVADESANCKSLHTLFCK 73
Db 45 IGVEHAKALMALPESQSLKCPHHEQVQVRVNVMDIADLCSRGAKHGDGCKSVMTILNE 164
QY 74 LCTVATLRETYGEMADCCAKQEPBERNECFLOHKKDNP-NLPLRVPEVDVMTAFPHDNE 132
Db 105 ICKTPEIPEKYPFHEGGCKKEDPERHKCFIEHKSTDPKERTYVVKPSPEQICKDHAE 164
QY 133 TFLKXYLYEIAARRHPYFAPELLTFKRYKAAFTCCQAADKAAACLLPKLDELHDEGRK 191
Db 165 EFLGHYIHKVASSHTTMYPPAILSFTLHPDGIIVSHCCKDEATVGCCLSEKMPAKHEVEH 224
QY 192 SSARQRLKASLQKGERAFKAVARLQRFPAKFAEVSKLVDTLTKVHTTCCCHGDL 251
Db 225 VCAYQKNCVILQNFERNALRASKAHACSKFPHASPVORLTGDIVHLHQTCCGDMY 284
QY 252 ECADRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDENPADLPSLAADPVE 311
Db 285 ACMAERMKLTQTQCEK-----KKCEKXPVLERSECIIVLPNDEKPADLSPEVRYFD 336
QY 312 SKOVCKNYBAKOVFLGMLFELYEARRRHPDYSVLLLRALAKTYETTTLEKCCAAADPHECYA 371

```



```

Db 455 STCCQLSEKXSCGSGMAD:FIGH:CIRNEASPVNSG:SHCCNSYSNRR:CTISFLAD 514
Qy 495 ETVVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKKHFKATKEQLKAWMDFAAF 554
Db 515 ETVAPPPFSEDKTIFPKDLCOAQKALQTMKQELLNLVKQKPELTTEEQLAAVTADFSGL 574
Qy 555 VEKCKADDKETCFABEGKLVAAASQAALGL 585
Db 575 LEKCKAQDOEVCTFEBGPKL:SKTRDALGV 605

RESULT 12
Q8JIA7
ID Q8JIA7 PRELIMINARY; PRT; 400 AA.
AC Q8JIA7
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE A/B over-sized serum albumin (Fragment).
CS Sphenodon puccatus (Bacteria) (Tuatara).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
OX NCBI_TaxID=8508;
RN [1]
RP SEQUENCE FROM N.A.
RA Metcalf V J., Brennan S.C., George P.M., Chambers G.K.;
R: "Partial mRNA sequence for tuatara A/B serum albumin."
RJ Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375973; AA046106.1; -.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
FT NON_TER 1
FT SIGNAL 1
SQ SEQUENCE 400 AA; 45715 MW; 8DE20609657CF753 CRC64;

Query Match 33.7%; Score 1045; DB 13; Length 400;
Best Local Similarity 47.7%; Pred. No. 5.1e-74;
Matches 186; Conservative 80; Mismatches 124; Indels 0; Gaps 0;

Qy 196 QRLKASLQKGFGERAFKAMAVARLSQRFPAEVSFKLVDTLKYHTSCCHGELLECAD 255
Db 3 EKHSGLXSGFERAFQANKLVLSQKFKAPFEIHKVVLATKLQKCGHGMIECLD 62
Qy 256 DRADLAKYICENQDSISSK:KECEKELLESKSHCIAEVDENPDLPLSLADDFVESKDY 315
Db 63 DRVEMAYICSKQAVFESKIKCCEKPIVQREYCIQACLDCEKPAJLPSIAGGYIESTEV 122
Qy 316 CKNYAEAKQVFLGMFLYEYARRHPDYSWLLRLAKTYETTLKCCAAADPHECYAKVFD 375
Db 123 CKHYESKQVFLAHFVYEYERRHPERSQMLLTGKGYQDTLXCKCTNPECYKAGE 182
Qy 376 EFKPLVEEPQNLIKONCELFEQGEYKFNALLVRYTKVPQVSTPTLVEVSANLKGVS 435
Db 183 ELARHIQESGELKTHCSFYTQCKDPQKMWLVRYTKVPQVPAELIEISKLTGVGV 242
Qy 436 KCGHPEAKMPCADYLSVNLQCVLHEKTPVSDRYV:KCCTESLVNRRPCEFALEVDE 495
Db 243 KCCPLSDKRLSCSEKHLSMVLFICRQHEASPVNNHVTTCCTDSYSEMRPCTFLGVGD 302
Qy 496 TYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKKHFKATKEQLKAWMDFAAFV 555
Db 303 SYVPEFCPTFLFDEQLCTAPEARLKKGLTLVKKLQKPOIEDEQLKLVDTWAVE 362
Qy 556 EKCKADDDKTCFABEGKLVAAASQAALGL 585
Db 363 EKCCQAENKQCEFTSEGEKLTQEGKALLGV 392

RESULT 13

```

```

Q8YGH6
ID Q8YGH6 PRELIMINARY; PRT; 603 AA.
AC Q8YGH6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Rana shqiperica.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.
OX NCBI_TaxID=44326;
RN [1]
RP SEQUENCE FROM N.A.
RA Jizzell T., Hotz H.;
RC STRAIN=Bushat; TISSUE=Liver;
RT "Albumin cDNA sequence of Rana shqiperica: evolutionary changes in
RL frog albumins."
DR EMBL; U40452; RAD09358.1; -.
DR HSSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Signal.
FT NON_TER 1
FT SIGNAL 23
FT CHAIN 24 603 SERUM ALBUMIN.
SQ SEQUENCE 603 AA; 69293 MW; 340D3723FA010C99 CRC64;

Query Match 30.8%; Score 955; DB 13; Length 603;
Best Local Similarity 33.6%; Pred. No. 9.9e-67;
Matches 190; Conservative 119; Mismatches 241; Indels 16; Gaps 3;

Qy 12 KDLGEENFKALVLIAPQYLLQCGPFEDHVKLVNVEVTEFAKTCVADESAENCDKSLHTLFG 71
Db 37 KAVGKPAVEKLVLMVAQDPFEKCSLDEHLKVQAKITAEAVDNCCKHPPEAECKKPALEYH 96
Qy 72 DKLTATLRETTGEMADCCAKOEPERNBCFLQHKDQNPMLPLVR-----REVDMCT 125
Db 97 DVCKEEDIDQLYPMWITTECGKAEARTKCFEHE-----VRVEYKFNIEESCK 148
Qy 126 AFHNEETFLKKYLYETARHPYFYAPELFFPAKRYKAAFTCCQADKAACLLPKLDEL 185
Db 149 EKHCHPGRAFYSYLSNTAKRHSKLYPPAVLGFAIQYNEITTECCAAEDKAKCGERMPQV 208
Qy 186 RDEKASSAKQRLKASLQKGFGERAFKAMAVARLSQRFPAEVSFKLVDTLKYHTEC 245
Db 209 KKLNYLEDKHKQCRVLKSFPERVSQALTLVQVSQRFQNAKYDDVEKVTFIEAHLNEDC 268
Qy 246 CHGDLLECCADRADLAKYICENQDSISSK:KECEKELLESKSHCIAEVDENPDLPLSL 305
Db 269 CKGDAVSCMTIERMETHICLAKELKSSKLSDCCAKGVLERTFCILALPNEE--PDLPIE 326
Qy 306 AADFEVSQKVCNVAEAKDVFGLGMFLYEYARRHPDYSWLLRLAKTYETTLKCCAAAD 365
Db 327 LKEYEEDHVCENYQDKRYLAHFTHDYSRSHQESSPQSLVSRGFEMLLKCCASAN 386
Qy 366 PHECYAKVDFEKPFLVEEPQNLIKONCELFEQGEYKFNALLVRYTKVPQVSTPTLVE 425
Db 387 SAELCKADAPKLLLEAALNEEIEISKQNCGALEKGLGNDFYIQLLVRYFGKMPQVTAOTLVE 446
Qy 426 VSRNLGKVSCKCHPEAKMPCADYLSVNLQCVLHEKTPVSDRYV:KCCTESLVNRR 485
Db 447 LTGRMAKIGVYCCGLPDNKKQPCAEKLDILGEMCEREKKTFINDNVHHCVDVSNRR 506
Qy 486 PCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKOTALVELVKKHFKATKEQLK 545
Db 507 PCFTKLGPIYANEPVWDESKLHFTADMCKGSADDQLKTKLVLLVVEFLQKPCGKEKLT 566

```

```
QY 546 AVMDZFAAFVEKCKKADDKETCPAAE 57:
Db 567 EVIESFKTVECCAAENQACFDK 592

RESULT 14
Q91134 PRELIMINARY; PRT; 614 AA.
ID Q91134
AC Q91134
DT 01-NOV-1996 (TREMBSrel. 01, Created)
DT 01-NOV-1996 (TREMBSrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBSrel. 21, Last annotation update)
DE Cobra serum albumin.
OS Naja naja (Indian cobra).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scieroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Naja.
OX NCBI_TaxID=35670;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=96145734; PubMed=856:913;
RA Wang X., Hansen H., Havsteen B.;
R "Evidence of the coevolution of snake toxin and its enogenous
RT antitoxin. Cloning, sequence and expression of a serum albumin cDNA of
RL the chinese cobra."
RL Bio. Chem. Hoppe-Seyler 376:545-553(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93343893; PubMed=8343135;
RA Shao J., Sher H., Havsteen B.;
R "Purification, characterization and binding interactions of the
RT Chinese-cobra (Naja naja atra) serum antitoxic protein CSAP."
RL Biochem. J. 293:559-566(1993).
DR EXBL; X78598; CAA55333.1; -.
DR HSSP; PC2768; 1E7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 614 AA; 69798 MW; 3DB2D3CC4BD8CBFD CRC64;

Query Match 29.9%; Score 928.5; DB 13; Length 614;
Best Local Similarity 32.8%; Pred. No. 1.2e-64;
Matches 184; Conservative 119; Mismatches 255; Indels 3; Gaps 3;

QY 18 NFKALVIAFAQYLQOCFFEDHVKLVNEVTEBFKTCVADE-SAEKCDKSLHTLFGDXLCT 76
Db 45 DFGA:TLTLVTPNATLEDLKLKSABIIELHKKCVASEFSDPCTKPLGIVFLDVLCH 104

QY 77 VATTRETYGEMADCAKOEPRNCFLOKDCNP-NLPRLVRPEVDVMCTAFHNEETFL 135
Db 105 NEEFSENKYG-INDCCAKADPRNECVLSHKTSSTSTISPPVHPNAEEACQAFQNDRESVL 163

QY 136 XKYLVEIARRRHYFVAPLELLFAKRYKAAFTCCQAAKCAKCLPKLDELRCBGKASSAK 195
Db 164 AQYFELSRRTALSVVILSESTYTKILTECCAEADKACIHEKATEAKKFRFVME 223

QY 196 QRLKASLOKGERAFKAWAVARISQRPKAEFAEVSXLVTDLTJKVHTTECGHGLLECAD 255
Db 224 QEYTCYNLKKYCKKLYALKFTIETHEKFNAKLETITIGIAEFVHHVEEICMGDSVVLV 283

QY 256 DRADLAKYICENQDSISKLKECCCKPLLEKSHCHTAIVENDEMPADLPSLAAPVESKDV 315
Db 284 DRAALSQYVCEHKAIDSSNVGHCCEKPLVRPNCLATLANDARSPCLPPPEEIIKETE 343

QY 316 CKNYAEAKDVLGYELVEYARRHPDYSVVLJLRUAKTYETTTLEKCCAAADPHECYAKVFD 375
Db 344 CTTYTEQRENYKESFLTJTRNHPELSKLIDLEILDYKYLEKLEBCCSEHHVQCLHGEQ 403
```

```
QY 376 EFKPLVEBPONLIKONCELFQGLGEYKFNALLVRYTKKVPQVSTPTLVEVSRNLGKVG 435
Db 404 VFKLITKINEVWKSNCDSYKELGDYFTTNEFLVKYSRMMPQAPTSTFLIELTERKVGKVAE 463

QY 436 KCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTESLVNRRPCFSALEVDE 495
Db 464 KCCNLDSNHQVSCALENTDKYMGSGICKYHNRHFINDQICHCCNSFTSRWECISNLGPDL 523

QY 496 TVVPKEFNAETFTFHAD:CTLSEKERQIKKOTALVELVKHKPKATKECLKAVMDFAAPV 555
Db 524 SFVPTFPNPKTMDNPEKJCSTSEOTVQSKKGLSELVKSKFNISEBELAATILTFRIQ 583

QY 556 EKCKKADDKETCPAEEGKKLV 576
Db 584 KLCCEAKKKECFDKKQEMV 604

RESULT 15
Q9JIA8 PRELIMINARY; PRT; 406 AA.
ID Q9JIA8
AC Q9JIA8;
DT 01-OCT-2002 (TREMBSrel. 22, Created)
DT 01-OCT-2002 (TREMBSrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBSrel. 23, Last annotation update)
DE Serum albumin (Fragment).
OS Haploactylus maculatus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scieroglossa; Gekkota; Gekkonidae;
OC Haploactylus
CX NCBI_TaxID=150461;
RN [1]
RP SEQUENCE FROM N.A.
RA Metcalf V.J., Brennan S.C., George P.M., Chambers G.K.;
R "Partial mRNA sequence for New Zealand gekko albumin."
RT Submitted (MAY-2001) to the EXBL/GenBank/DBS databases.
DR EMBL; AF375972; AAM46105.1; -.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
FT NON TER
SQ SEQUENCE 406 AA; 46679 MW; 03CD838FA3272E10 CRC64;

Query Match 28.6%; Score 888; DB 13; Length 406;
Best Local Similarity 42.0%; Pred. No. 1.1e-61;
Matches 173; Conservative 74; Mismatches 157; Indels 8; Gaps 2;

QY 172 ADKAACLLPKLDELDEGKASSAKQRLKCAISLOKGERAFKAWAVARISQRPKAEFAEV 231
Db 2 AEKSPATKKL-----KEATLMKQNCFLVKFGPKELHTWKPAQLAKQKPKADRFVL 54

QY 232 SKLVTDLTJKVHTTECGHGLLECADRADLAKYICENQDSISKLKECCCKPLLEKSHCHTA 291
Db 55 HNITHDIVHVKERCGRGDTLESYLDQVQISRVRCGHQDLFSPKVKKCCDD:LLHRSPECLV 114

QY 292 EVENDEMPADLPSLAADFVESKDVCKVVAEKDVFSGMFLYEVARRHPDYSVVLLLRLLAK 351
Db 115 AVENDEPPADLSPTVREFVDNKEVQCFPAENHDDHLERFVYEGRHHQDFSPOLLQORJCK 174

QY 352 TYETTTLEKCCAAADPHECYAKVDFEKLPLVEBPONLIKONCELFQGLGEYKFNALLVRY 411
Db 175 GYHDLLEKCCLEAPACLLGEPLLKHVADTLAVISTNCDLYAKLGYKFNKMLLARY 234

QY 412 TKKVPQVSTFTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSD 471
Db 235 TKKAPOLSTFEUSQYAEQRLTVANKCCQGLGDEKKLICGSIYTDVJGNICJHHQHPINP 294

QY 472 RVTKCTESLVNRRPCFSALEVDETVVPKEFNAETFTFHAD:CTLSEKERQIKKOTALVE 531
Db 295 RICQCCSNFYAFRECFSSIEVDEGVVPAPCAPELTFHEDLCT-ANTDVEKKKQKLLVD 553
```

OY 532 LVKHKFKATKEQLKAVMDFAAFVEKCKCCKADDKETCFPAEEGKKLVAASQAA 553
Db 354 VVKCRPAITHEQLKAVITDFYGVWEKCCCHGENHEACFLAEGFCLVQRTCAA 455

Search completed: October 27, 2003, 15:21:05
Job time : 108 secs